

BEST AVAILABLE COPY

STIC-Biotech/ChemLib

173.922

Mg

From: CRFE Marvich, Maria
Sent: Monday, December 12, 2005 7:33 AM
To: STIC-Biotech/ChemLib
Subject: RE:

10/648361

Please search SEQ ID NO:2 (not interference) 187 amino acids.

Thank you

-----Original Message-----
From: STIC-Biotech/ChemLib
Sent: Monday, December 12, 2005 7:32 AM
To: Marvich, Maria
Subject: RE:

I need your Serial number for this request. maude

-----Original Message-----
From: Marvich, Maria
Sent: Sunday, December 11, 2005 10:12 AM
To: STIC-Biotech/ChemLib
Subject:

Please search SEQ ID NO:2 (not interference) 187 amino acids.

Maria Bonovich Marvich, PhD
United States Patent and Trademark Office
Remsen 2B84
AU 1633
Mail Box 2C70
571-272-0774

Searcher: _____
Searcher Phone: _____
Date Searcher Picked up: _____
Date completed: _____
Searcher Prep Time: _____
Online Time: _____

Type of Search
NA# _____ AA# _____
S/L: _____ Oligomer: _____
Encode/Transl: _____
Structure #: _____ Text: _____
Inventor: _____ Litigation: _____

Vendors and cost where applicable
STN: _____
DIALOG: _____
QUESTEL/ORBIT: _____
LEXIS/NEXIS: _____
SEQUENCE SYSTEM: _____
WWW/Internet: _____
Other (Specify): _____

Searcher: Beverly e 2528
Terminal time: _____
Elapsed time: _____
CPU time: _____
Total time: _____
Number of Searches: _____
Number of Databases: _____

Type of Search
____ STIC
____ CM-1
____ Pre-S
____ N.A. Sequence
____ A.A. Sequence
____ Structure
____ Bibliographic

Vendors
____ IG
____ STN
____ Dialog
____ APS
____ Geninfo
____ SDC
____ DARC/Questel
____ Other ☒ CGN

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 13, 2005, 02:56:16 ; Search time 230 seconds
(without alignments)
573.625 Million cell updates/sec

Title: US-10-648-361-2

Perfect score: 983

Sequence: 1 MWCASPVAVVAFCAGLLVSH.....VNLQELAPKYAATHPKTEL 187

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 216643 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt_05.80.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	768	78.1	187	2	Q8CC36 mus musculus
2	538	54.7	179	1	ARMET_HUMAN
3	532	54.1	179	1	ARMET_MOUSE
4	530	53.9	180	2	Q3ZM4_XENLA
5	530	53.9	201	2	Q7ZYC7_XENLA
6	505	51.4	165	2	Q80ZP8_MOUSE
7	430	43.7	126	2	Q7SZ63_XENLA
8	399.5	40.6	172	2	Q7QD98_ANOGA
9	393	40.0	173	1	ARMET_DROME
10	364.5	37.1	168	1	ARMET_CABEL
11	356	36.2	169	2	Q61B22_CAEBR
12	344	35.0	147	2	Q6B882_9ACAR
13	334	34.0	152	2	Q4T0D1_TETNG
14	314.5	32.0	184	2	Q5DHK5_SCHJA
15	95.5	9.7	1558	2	Q4SEW9_TETNG
16	93.5	9.5	540	2	Q4PW48_METMA
17	93.5	9.5	1071	2	Q96V10_PNECA
18	92	9.4	283	2	Q7VF6E_HELHP
19	92	9.4	294	2	Q95X03_NAEFO
20	92	9.4	307	2	Q9BKM2_NAEFO
21	92	9.4	853	2	Q4P3R3_USTMA
22	91	9.3	481	2	Q82T66_NITEU
23	90.5	9.2	1026	2	Q74669_PNEJI
24	89.5	9.1	299	2	Q5CU74_CRYPV
25	89.5	9.1	299	2	Q5CE54_CRYHO
26	89.5	9.1	309	2	Q810T8_MOUSE
27	89.5	9.1	347	2	Q5BH60_EMENI
28	89.5	9.1	588	2	Q3R46_SHIFL
29	89.5	9.1	610	1	UVRC_ECO57
30	89.5	9.1	610	1	UVRC_ECOL6
31	89.5	9.1	610	1	UVRC_ECOLI

32	89.5	9.1	775	1	TNAP3_MOUSE	Q60769 mus musculus
33	89.5	9.1	775	2	Q8C2N2_MOUSE	Q8c2n2 mus musculus
34	89.5	9.1	775	2	Q7TOD1_MOUSE	Q7tqd1 mus musculus
35	89.5	9.1	1028	2	Q74668_PNEJI	Q74668 pneumocysti
36	89.5	9.1	1119	2	Q5GTB1_WOLTR	Q5gtb1 wolbachia s
37	89	9.1	39	1	ARMET_BOVIN	P80513 bos taurus
38	89	9.1	498	1	CP6B1_PAPPO	Q04552 papilio pol
39	89	9.1	563	1	DPO3X_BACSU	P09122 bacillus su
40	88.5	9.0	506	2	Q4NSP0_THERA	P09122 bacillus su
41	88.5	9.0	523	2	Q8TT47_METAC	Q4n5p0 theileria p
42	88.5	9.0	828	2	Q65503_ATHAT	Q65503 arabidopsis
43	88	9.0	217	2	Q6L4C3_SOLDE	Q614c3 solanum dem
44	88	9.0	953	2	Q81IH2_BACCR	Q81ih2 bacillus ce
45	88	9.0	1018	2	Q4PAL9_USTWA	Q4pal9 ustilago ma

ALIGNMENTS

RESULT 1
Q8CC36_MOUSE
ID Q8CC36_MOUSE PRELIMINARY; PRT; 187 AA.
AC Q8CC36;
DT 01-MAR-2003 (Tremblrel. 23, Created)
DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)
DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)
DE Mus musculus adult male diencephalon cDNA, RIKEN full-length enriched
DE library, clone:933014G23 product:hypothetical protein, full insert
DE sequence.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Diencephalon;
RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P., Hayashizaki Y.,
RT "High-efficiency full-length cDNA cloning."
RL Meth. Enzymol. 303:19-44(1999).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Diencephalon;
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
RA Schirai L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Momotarts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.,
RT "Functional annotation of a full-length mouse cDNA collection."
RL Nature 409:685-690(2001).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Diencephalon;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RA "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs."
RL Nature 420:563-573(2002).

```

RN [4]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Diencephalon;
RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630(2000).
RN [5]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Diencephalon;
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA Konno H., Akiyama J., Nishi K., Kitsuai T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yamanoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1757-1771(2000).
RN [6]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Diencephalon;
RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
RA Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,
RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
RA Katoh H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,
RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,
RA Saito K., Saiton H., Sakai K., Sakai K., Sakazume N., Sano H.,
RA Sasaki D., Shibata K., Shingawa A., Shiraki T., Sogabe Y., Tagami M.,
RA Tagawa A., Takahashi F., Takaku-Akashira S., Takeda Y., Tanaka T.,
RA Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR ENBL; AK034009; BAC28454.1; -; mRNA.
DR Ensembl; ENSMUSG0000039496; Mus musculus.
DR InterPro; IPR001545; Gly_hormoneB.
DR PROSITE; PS00261; GLYCO_HORMONE_BETA_1; UNKNOWN_1.
KW Hypothetical protein.
SQ
SEQUENCE 187 AA; 21031 MW; 24DFPE42026A7971 CRC64;

Query Match 78.1%; Score 768; DB 2; Length 187;
Best Local Similarity 77.5%; Pred. No. 4.4e-56;
Matches 145; Conservative 16; Mismatches 26; Indels 0; Gaps 0;

QY 1 MWCASPVAVVAFAGLLVSHPVLTQGOEAGRPGADCEVCCKEFLNRPYKSLIDRGVNFSL 60
DB 1 MRCISPTALVTFAGFCISNPLVLAQGLEAGVGRADCEVCCKEFLDRFYNLSLGRGIDFSA 60
QY 61 DTIEKELISFLDTKGNKRLCYLKGATKDAATKILSEVTRPMSVHPMPAKIKCEKLUKLD 120
DB 61 DTIEKELLNFCSDAKGNKRLCYLKGATTTAATKILGEVTRPMSVHIPAVKICEKLUKMD 120
QY 121 SQICELKYEXTLDLASVDLKPMAVELKQILHWSGEECRACAEKTDVYVNLIOELAPKYAA 180
DB 121 SQICELYKGGKLDLASVDLWKMVAELKQILRWGEECRACAEKSDVYVNLIRELAPKYVE 180
QY 181 THPKTEL 187
DB 181 IYQTEL 187

RESULT 2
ID ARMET HUMAN STANDARD; PRT; 179 AA.
AC P55145; O86U67; O95IS4;
DT 01-OCT-1996 (Rel. 34, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)

```

```

DE ARMET protein precursor (Arginine-rich protein).
GN Name=ARMET; Synonyms=ARP;
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE [MRNA], AND DISCUSSION OF A PUTATIVE CANCER
RP VARIANT.
RX MEDLINE=96211400; PubMed=8649854;
RA Shridhar V., Rivard S., Shridhar R., Mullins C., Bostick L., Sakr W.,
RA Grignon D., Miller O.J., Smith D.I.;
RT "A gene from human chromosomal band 3p21.1 encodes a highly conserved
RT arginine-rich protein and is mutated in renal cell carcinomas.";
RL Oncogene 12:1931-1939(1996).
RN [2]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RA Kalline N., Chen X., Rolfs A., Halleck A., Hines L., Eisenstein S.,
RA Koundinya M., Raphael J., Moreira D., Kelley T., LaBaer J., Lin Y.,
RA Phelan M., Farmer A.;
RT "Cloning of human full-length CDSs in BD Creator(TM) system donor
RT vector.";
RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
RN [3]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RC TISSUE=Eye;
RX MEDLINE=22388957; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altshuler S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heide F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Frange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalish D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [4]
RP PROTEIN SEQUENCE OF 22-36.
RC TISSUE=platelet;
RX MEDLINE=22608258; PubMed=12665801; DOI=10.1038/nbt810;
RA Gevaert K., Goethals M., Martens L., Van Damme J., Staes A.,
RA Thomas G.R., Vandekerckhove J.;
RT "Exploring proteomes and analyzing protein processing by mass
RT spectrometric identification of sorted N-terminal peptides.";
RL Nat. Biotechnol. 21:566-569(2003).
RN [5]
RP DISCUSSION OF A PUTATIVE CANCER VARIANT.
RX MEDLINE=97126232; PubMed=8971156;
RA Shridhar R., Shridhar V., Rivard S., Siegfried J.M.,
RA Pietraszkiewicz H., Ensley J., Pauley R., Grignon D., Sakr W.,
RA Miller O.J., Smith D.I.;
RT "Mutations in the arginine-rich protein gene, in lung, breast, and
RT prostate cancers, and in squamous cell carcinoma of the head and
RT neck.";
RL Cancer Res. 56:5576-5578(1996).
RN [6]
RP DISCUSSION OF PUTATIVE CANCER VARIANTS.
RX MEDLINE=97316783; PubMed=9174057; DOI=10.1038/sj.onc.1201054;
RA Shridhar V., Rivard S., Wang X., Shridhar R., Paisley C., Mullins C.,
RA Beirnat L., Dugan M., Sarkar P., Miller O.J., Vaitkevicius V.K.,
RA Smith D.I.;

```

"Mutations in the arginine-rich protein gene (ARP) in pancreatic cancer.";
 RT Oncogene 14:2213-2216(1997).
 RL
 CC -1- SUBCELLULAR LOCATION: Secreted (Potential).
 CC -1- SIMILARITY: Belongs to the ARMET family.
 CC -1- CAUTION: Was originally (Ref.1, Ref.5 and Ref.6) thought to be
 CC much longer and included an arginine-rich region thought to be the
 CC target of cancer-causing mutations. All these mutations are in
 CC what is now the 5' non-translated region of the mRNA and the gene.
 CC
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use as long as its content is in no way modified and this statement is not
 CC removed.
 CC
 CC EMBL; M83751; AB08753.1; ALT_INT; mRNA.
 CC EMBL; BT007110; AAP35774.1; ALT_INIT; mRNA.
 CC EMBL; BC007282; AAH07282.1; ALT_INIT; mRNA.
 CC OGP; P55145; -.
 CC Ensembl; ENSG00000145050; Homo sapiens.
 CC HGNC; HGNC:15461; ARMET.
 CC MIM; 601916; -.
 CC Direct protein sequencing; Signal.
 KW SIGNAL 1 21
 FT CHAIN 22 179 ARMET protein.
 FT CONFLICT 10 11 AL -> RV (in Ref. 1).
 FT CONFLICT 176 176 R -> P (in Ref. 1).
 SQ SEQUENCE 179 AA; 20257 MW; 4CCCCABF8208A73B CRC64;
 Query Match 54.7%; Score 538; DB 1; Length 179;
 Best Local Similarity 56.6%; Pred. No. 6.7e-37;
 Matches 107; Conservative 26; Mismatches 44; Indels 12; Gaps 4;
 QY 1 MWCASPVAVVAFAGLVSHPVLTQGEAGRGADCEVCCKEFLNRYKSLIDRGVNFSL 60
 DB 1 MWATQGLA-VALALSVPGRSAL-----RPG-DCEVCISYLGFRYQDLKDRDVTFS 50
 QY 61 DTTEKELISFCLDTKGNRLCYVLGATKDAATKILSEVTRPMSVHMPAMKICEKLKLD 120
 DB 51 ATTEELIKFCREARGENRLCYIGATDAAATKIINEVSKPLAHHPVKEKICELKKKD 110
 QY 121 SQICELKYKTDLASVDLRMRVAELKQILHSWGECRACAEKTDYVNLIOELAPKYA- 179
 DB 111 SQICELKYDQIDLSVDLKKLVKELKILDDGECCKGCAEKSDYIRKINELMPKYAP 170
 QY 180 -ATHPKTEL 187
 DB 171 KAASARTDL 179
 RESULT 3
 ID ARMET_MOUSE STANDARD; PRT; 179 AA.
 AC Q9CX15;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 10-MAY-2005 (Rel. 47, Last annotation update)
 DE ARMET protein precursor.
 GN Name=Armlet;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
 OC Muridae; Muridae; Murinae; Mus.
 OC NCBI_TaxID=10090;
 RN [1]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
 RC STRAIN=C57BL/6J; TISSUE=Embryonic head.
 RX MEDLINE=22354683; PubMed=12456851; DOI=10.1038/nature01266;
 RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
 RA Nakada I., Osato N., Saito K., Suzuki H., Yananaka I., Kiyosawa H.,
 RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,
 RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,

Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
 RA Blake J.A., Bradt D., Brusic V., Ciochia C., Corbani L.E., Cousins S.,
 RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
 RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
 RA Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
 RA Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,
 RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
 RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
 RA Nagashima T., Numata K., Okido T., Pavan W.J., Perlea G., Pesole G.,
 RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
 RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
 RA Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K.,
 RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
 RA Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
 RA Wilming L.G., Wyshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
 RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
 RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
 RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
 RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
 RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
 RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
 RA Birney E., Hayashizaki Y.;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs";
 RL Nature 420:563-573(2002).
 CC -1- SUBCELLULAR LOCATION: Secreted (Potential).
 CC -1- SIMILARITY: Belongs to the ARMET family.
 CC
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use as long as its content is in no way modified and this statement is not
 CC removed.
 CC
 CC EMBL; AK014338; BAB29281.1; -; mRNA.
 CC Ensembl; ENSMUSG0000032575; Mus musculus.
 CC MGI; MGI:1922090; Armlet.
 CC GO; GO:0005615; C:extracellular space; TAS.
 KW Signal.
 FT SIGNAL 1 21 Potential.
 FT CHAIN 22 179 ARMET protein.
 SQ SEQUENCE 179 AA; 20374 MW; E5BCBE8C033C1530 CRC64;
 Query Match 54.1%; Score 532; DB 1; Length 179;
 Best Local Similarity 56.1%; Pred. No. 2.1e-36;
 Matches 106; Conservative 27; Mismatches 44; Indels 12; Gaps 4;
 QY 1 MWCASPVAVVAFAGLVSHPVLTQGEAGRGADCEVCCKEFLNRYKSLIDRGVNFSL 60
 DB 1 MWATRGLA-VALALSVPDSRAL-----RPG-DCEVCISYLGFRYQDLKDRDVTFS 50
 QY 61 DTTEKELISFCLDTKGNRLCYVLGATKDAATKILSEVTRPMSVHMPAMKICEKLKLD 120
 DB 51 ATTEELIKFCREARGENRLCYIGATDAAATKIINEVSKPLAHHPVKEKICELKKKD 110
 QY 121 SQICELKYKTDLASVDLRMRVAELKQILHSWGECRACAEKTDYVNLIOELAPKYA- 179
 DB 111 SQICELKYDQIDLSVDLKKLVKELKILDDGECCKGCAEKSDYIRKINELMPKYAP 170
 QY 180 -ATHPKTEL 187
 DB 171 KAASARTDL 179
 RESULT 4
 ID Q63ZM4_XENLA PRELIMINARY; PRT; 180 AA.
 AC Q63ZM4;
 DT 25-OCT-2004 (TEMBLrel. 28, Created)
 DT 25-OCT-2004 (TEMBLrel. 28, Last sequence update)
 DT 25-OCT-2004 (TEMBLrel. 28, Last annotation update)
 DE Armlet protein.
 GN Name=Armlet;

OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC Xenopodinae; Xenopus; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Eye;
RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.;
RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT initiative.";
RL Dev. Dyn. 225:384-391(2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Eye;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Eye;
RA Klein S., Gerhard D.S.;
RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC082888; AAH82888.1; -; mRNA.
DR InterPro; IPR008139; SaposinB.
SQ SEQUENCE 180 AA; 20198 MW; 3448F4A2818615D1 CRC64;

Query Match 53.9%; Score 530; DB 2; Length 180;
Best Local Similarity 54.9%; Pred. No. 3.1e-36;
Matches 101; Conservative 33; Mismatches 42; Indels 8; Gaps 3;

QY 6 PVAVAFACGLLVSHPVLTQGOBAGCPGACDCEVCKEFLNRFYKSLIDRGVNFSLDTIEK 65
Db 3 PLALLT-VTGIMVLLP-----SDAGALKAGDCEVCISFSLRFSYQSLKRVKPEKPIVEK 56

QY 66 ELISFCLDTKGKENCRLCYVIGATKDAATKILSVTRPMSVHMPAMKICCKLKKLDSQICE 125
Db 57 ELLKTCNDARGKENRLCYIGATSDAATKITNEVSKPLSHIPEKICEKLLKKKQDQICE 116

QY 126 LKYEKTLDLASVLRKMRVAELKQILHSWGECRACAEKTDYVNLQELAPKYA--ATHP 183
Db 117 LKYDKQIDLSTVDLKKLVKELKILLDDWGESCCKGCAEKSDFKIRKINELMPKYPAPNAANA 176

QY 184 KTEL 187
Db 177 RTDL 180

RESULT 5
Q7ZYC7_XENLA
ID Q7ZYC7_XENLA PRELIMINARY; PRT; 201 AA.
AC Q7ZYC7;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)

DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Armet protein (Fragment).
GN Name=Armet;
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC Xenopodinae; Xenopus; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Embryo;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Embryo;
RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.;
RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT initiative.";
RL Dev. Dyn. 225:384-391(2002).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Embryo;
RA Klein S., Strausberg R.;
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC043846; AAH43846.1; -; mRNA.
DR InterPro; IPR008139; SaposinB.
FT NON TER 1 1
SQ SEQUENCE 201 AA; 22414 MW; 574084EB920E1518 CRC64;

Query Match 53.9%; Score 530; DB 2; Length 201;
Best Local Similarity 54.9%; Pred. No. 3.5e-36;
Matches 101; Conservative 33; Mismatches 42; Indels 8; Gaps 3;

QY 6 PVAVAFACGLLVSHPVLTQGOBAGCPGACDCEVCKEFLNRFYKSLIDRGVNFSLDTIEK 65
Db 24 PLALLT-VTGIMVLLP-----SDAGALKAGDCEVCISFSLRFSYQSLKRVKPEKPIVEK 77

QY 66 ELISFCLDTKGKENCRLCYVIGATKDAATKILSVTRPMSVHMPAMKICCKLKKLDSQICE 125
Db 78 ELLKTCNDARGKENRLCYIGATSDAATKITNEVSKPLSHIPEKICEKLLKKKQDQICE 137

QY 126 LKYEKTLDLASVLRKMRVAELKQILHSWGECRACAEKTDYVNLQELAPKYA--ATHP 183
Db 138 LKYDKQIDLSTVDLKKLVKELKILLDDWGESCCKGCAEKSDFKIRKINELMPKYPAPNAANA 197

QY 184 KTEL 187
Db 198 RTDL 201

RESULT	7
Q7SZ63_XENLA	
ID	Q7SZ63_XENLA PRELIMINARY;
AC	Q7SZ63; PRT; 126 AA.
DT	01-OCT-2003 (TrEMBLrel. 25, Created)
DD	01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT	01-OCT-2003 (TrEMBLrel. 25, Last annotation update)

```

DE ENSANGP0000018242 (Fragment).
GN ORFNames=ENSANGG00000015753;
OS Anopheles gambiae str. PEST.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Culicidae;
OC Anophelinae; Anopheles.
OX NCBI_TaxID=180454;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PEST;
RG The Anopheles gambiae Sequence Committee;
RT "Anopheles gambiae re-annotation.";
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
[2]
RN NUCLEOTIDE SEQUENCE.
RC STRAIN=PEST;
RG The Anopheles gambiae Sequence Committee;
RL Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAAB01008859; EAA07858.2; -; Genomic DNA.
FT NON_TER
SQ
SEQUENCE 172 AA; 19874 MW; 05B1E3519191B28 CRC64;

Query Match 40.6%; Score 399.5; DB 2; Length 172;
Best Local Similarity 46.4%; Pred. No. 2.3e-25;
Matches 85; Conservative 30; Mismatches 47; Indels 21; Gaps 5;

QY 7 VAVAFACGILLVHPVLTQEQAGRGADCEYCKEFLNRYKSLIDRGVNFSLDT--IE 64
DB 9 VCILFP---LLPHSTALREG-----DCEVCVKTVNTFMTLSD---ETKKDKRIE 53
QY 65 KELISCLDTKGENKLCYVLTGATKDAATKILSEVTRPMSVHPAMKICEKLKLDQIC 124
DB 54 DEPRAFCKSKNKEQRFYVLTGVEDSATGILGELSKPISWSMPAEKICEKLKLDQIC 113
QY 125 ELKYEKTLDLASVDLRKMRVAELKQILHSGESCACAEKTDVNLIOELAPKYAATHPK 184
DB 114 DLRYDQIDVNDLKKLRDULKILSDWDEDCGCKLEKTFDKRIEKLKHV----K 169
QY 185 TEL 187
DB 170 TEL 172

RESULT 9
ARMET DROME STANDARD; PRT; 173 AA.
AC QXZ63; Q9VFW4;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE ARMET-like protein precursor.
GN Names=ARP-like; ORFNames=CQ7013;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Oregon-R;
RX MEDLINE=20062184; PubMed=10597048;
RA Geo J.H., Ann Y., Park O.K., Park W.J.;
RT "Selection of Drosophila genes encoding secreted and membrane
RT proteins.";
RL Mol. Cells 9:564-568(1999).
RN [2]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=Berkley;
RX MEDLINE=20196006; PubMed=10711132; DOI=10.1126/science.287.5461.2185;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,

```

```

RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter B.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Balley R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Gerhart W.M., Glasser K.,
RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glöckle A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
RA Liu X., Mattel B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siddle-Klanos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasarman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-P., Zaveri J.S., Zhan M.C., Zhang G., Zhao X., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
RN [3]
RP GENOME REANNOTATION.
RX MEDLINE=22426069; PubMed=12537572;
RA Misra S., Crosby M.J., Mungall C.J., Matthews B.B., Campbell K.S.,
RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochuk S.E.,
RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
RA Bettencourt B.R., Celniker S.E., de Grey A.D.N.J., Drysdale R.A.,
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA Lewis S.E.;
RT "Annotation of the Drosophila melanogaster euchromatic genome: a
RT systematic review.";
RL Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).
RN [4]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RC STRAIN=Berkley; TISSUE=Embryo;
RX MEDLINE=22426066; PubMed=12537569;
RA Stapleton M., Carlson J.W., Brokstein P., Yu C., Champe M.,
RA George R.A., Guarin H., Kronmiller B., Pacleb J.M., Park S., Wan K.H.,
RA Rubin G.M., Celniker S.E.; cDNA resource.";
RT "A Drosophila full-length cDNA resource.";
RL Genome Biol. 3:RESEARCH0080.1-RESEARCH0080.8(2002).
CC -!- INTERACTION:
CC Q8T390:endoA; NbExp=1; IntAct=EBI-135963; EBI-150782;
CC -!- SUBCELLULAR LOCATION: Secreted (Potential).
CC -!- SIMILARITY: Belongs to the ARMET family.
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL; AF132912; AAC32615.1; -; mRNA.
CC EMBL; AS003713; AAP55303.1; -; Genomic DNA.
CC EMBL; AY061080; AAL28628.1; -; mRNA.
CC -----
DR
DR
DR

```


RESULT 11	
Q61B22 CAEBR	
ID Q61B22 CAEBR PRELIMINARY;	PRT; 169 AA.
AC Q61B22;	
DT 25-OCT-2004	(TREMBLrel. 28, Created)
DT 25-OCT-2004	(TREMBLrel. 28, Last sequence update)
DT 25-OCT-2004	(TREMBLrel. 28, Last annotation update)
DE	Hypothetical protein CBG13529.

OX	NCBI_taxID=6238;
RN	[1]
RP	NUCLEOTIDE SEQUENCE.
RG	The C. briggsae Sequencing Consortium;
RL	Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
CC	- CAUTION: The sequence shown here is derived from an
CC	EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC	preliminary data.
DR	EMBL; CAAC0100064; CAB67929.1; -- Genomic_DNA.
DR	InterPro; IPR000886; ER target S.
DR	ProSITE; PS00014; ER_TARGET; UNKNOWN_1.
KW	Hypothetical protein.
SQ	SEQUENCE 169 AA; 19180 MW; 4B370E34FCA4FB45 CRC64;

Db	82	NEWTKPLSWNP	TDKVC	KDKL	KS	DAQI	CE	KY	DKPL	DWKT	IDL	KM	R	KEL	K	NIL	GEW	141
Qy	156	EERACAEKTD	YVNL	IQEL	AP	KYA	A	TH	PK	TEL	187							
Db	142	EACGCTEKSE	FTKRI	EEL	AP	KV	----	KO	EL	169								

Q95802, 2004 (TReMBLrel. 28, Created)
 23-OCT-2004 (TReMBLrel. 28, Last sequence update)
 25-OCT-2004 (TReMBLrel. 28, Last annotation update)
 25-OCT-2004 (TReMBLrel. 28, Last annotation update)
 DE ARMEY-like protein (Fragment).
 OS Ixodes pacificus (western blacklegged tick).
 OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;
 OC Parasitiformes; Ixodida; Ixodoidea; Ixodidae; Ixodes.
 OX NCBI TaxID:29930;

KN
RP
RC
RA
RT

[1]
NUCLEOTIDE SEQUENCE.
TISSUE=Salivary gland;
Ribeiro J.M.C.;
"An insight into the transcriptome of the adult

RT "An insight into the transcriptome of the salivary glands of the adult

```

RT female tick, Ixodes pacificus.";
RL Submitted (AUG-2004) to the EMBL/GenBank/DBJ databases.
RN [2]
RN NUCLEOTIDE SEQUENCE.
RC TISSUE=Salivary gland;
RA Francischetti I.M., Lane R.S., Pham V.M., Ribeiro J.M.C.;
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY674265; AAT92198.1; -; mRNA.
DR InterPro: IPR000986; ER_caret_S.
DR PROSITE: PS00014; ER_TARGET; UNKNOWN_1.
FT NON_TER
SQ SEQUENCE 147 AA; 16750 MW; AD1E596A1596E9A6 CRC64;

Query Match 35.0%; Score 344; DB 2; Length 147;
Best Local Similarity 49.3%; Pred. No. 8.5e-21;
Matches 62; Conservative 26; Mismatches 34; Indels 4; Gaps 1;

QY 62 TIBKELISFCLDTKGENRLCYVIGATKDAATKILSEVTRPMSVHMPAMKICEKLLKJDS 121
   ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 26 SVEAGLFECSTKSGPEHRECYVVGLEESATKIVNELTKPFGWGPALKVCEKLVAKDS 85

QY 122 QICELKYKTVLDIASVDLRMRVAELKQILHSGWCECRACAEKTDYVNLIOELAPKYAT 181
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 86 QICDLKPKVIDLTKVLDLKKLVKDLKILSDWDERCEGCVKTDVFKRIEEL----KTV 141

QY 182 HPKTEL 187
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 142 HMREEL 147

RESULT 13
Q4T0D1_TETNG
ID Q4T0D1_TETNG PRELIMINARY; PRT; 152 AA.
AC Q4T0D1_TETNG
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DE Chromosome undetermined SCAFI1227, whole genome shotgun sequence.
GN ORFNames=GSTENG000940001;
OS Tetraodon nigroviridis (Green puffer).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC Tetraodontoidea; Tetraodontidae; Tetraodon.
OX NCBI_TaxID=99883;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Jaillon O., Aury J.M., Brunet F., Petit J.L., Stange-Thomann N.,
RA Mauceli E., Bouneau L., Fischer C., Ozouf-Costaz C., Bernot A.,
RA Nicaud S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segurens B.,
RA Dasilva C., Salanoubat M., Levy M., Boudet N., Castellano S.,
RA Anthouard V., Jubin C., Castelli V., Katinka M., Vacherie B.,
RA Blemont C., Skalli Z., Cattolico L., Poulain J., De Berardinis V.,
RA Parra G., Lardier G., Chappelle P., Coutanceau J.P., Gouzy J.,
RA Kellie M., Wolff J.N., Guigo R., Zody M.C., Mesirov J.,
RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
RA Laudet V., Schachter V., Quetier F., Saurin W., Scarpelli C.,
RA Wincker P., Lander E.S., Weissbach J., Roest Crolius H.;
RT "Genome duplication in the teleost fish Tetraodon nigroviridis reveals
the early vertebrate proto-karyotype.";
RL Nature 431:946-957(2004).
RN [2]
RN NUCLEOTIDE SEQUENCE.
RG Genoscope; Whitehead Institute Centre for Genome Research;
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL: CRAE01011227; CAFE93651.1; -; Genomic DNA.
SQ SEQUENCE 152 AA; 17000 MW; CD62AA304833FB15 CRC64;

Query Match 34.0%; Score 334; DB 2; Length 152;

```

```

Best Local Similarity 49.7%; Pred. No. 6e-20;
Matches 73; Conservative 19; Mismatches 53; Indels 2; Gaps 2;

QY 9 VVAFCAGLLVSHPVLTQGOEAGRGADCEVCKEFLNRPYKSLIDRGVNFSLDTTKEKLI 68
   ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 1 MLAVC-GLSVALALTLPGPAEALKDGECEVCVTFGLGRFYDLKENDVXFNDVEIEKAML 59

QY 69 SFCLDTKGENRLCYVIGATKDAATKILSEVTRPMSVHMPAMKICEKLLKJDSQICELKY 128
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 60 RSKDAKGKRNPCFYIYGATSDAATKMINSEVRPMSSHVPEKICEKLLKJDSQICELKY 119

QY 129 EKTDLIASVDLRMRVAELKQI-LHSW 154
   ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 120 GEEEDLATGSCRNNCPCLPQTNSW 146

RESULT 14
Q5DHK5_SCHJA
ID Q5DHK5_SCHJA PRELIMINARY; PRT; 184 AA.
AC Q5DHK5_SCHJA
DT 10-MAY-2005 (TrEMBLrel. 30, Created)
DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DE Hypothetical protein.
OS Schistosoma japonicum (Blood fluke).
OC Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea; Strigeidida;
OC Schistosomatoidea; Schistosomatidae; Schistosoma.
OX NCBI_TaxID=6182;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Han Z.;
RT "The full-length cDNA sequences of Schistosoma japonicum genes.";
RL Submitted (NOV-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY812969; AAW24701.1; -; mRNA.
KW Hypothetical protein.
SQ SEQUENCE 184 AA; 21068 MW; 2C33E6048DE47D8C CRC64;

Query Match 32.0%; Score 314.5; DB 2; Length 184;
Best Local Similarity 40.3%; Pred. No. 3.1e-18;
Matches 58; Conservative 32; Mismatches 53; Indels 1; Gaps 1;

QY 35 ADCEVCCKEFLNRPYKSLIDRGVNFSLDTTKEKLIISFCLDTKGENRLCYVIGATKDAATK 94
   ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 28 SNCEVCIKPMGSPFIQSLPSDVD-SSDNIKQAFMKKCESSVGKDNDFCYVVGGLKTSAA 86

QY 95 ILSEVTRPMSVHMPAMKICEKLLKJDSQICELKYKTVLDIASVDLRMRVAELKQILHWS 154
   ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 87 TVNRLVDPIKWKMPVEKVCQKLFELDSQICDLRYEKLIDFKPEFDEKSKVKDLKKIMAK 146

QY 155 GEECRACAEKTDYVNLIOELAPKY 178
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 147 GLECRGCTEKKDPISLIKSNMHKH 170

RESULT 15
Q4SEM9_TETNG
ID Q4SEM9_TETNG PRELIMINARY; PRT; 1558 AA.
AC Q4SEM9_TETNG
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DE Chromosome undetermined SCAFI4615, whole genome shotgun sequence.
GN ORFNames=GSTENG00019474001;
OS Tetraodon nigroviridis (Green puffer).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC Tetraodontoidea; Tetraodontidae; Tetraodon.
OX NCBI_TaxID=99883;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Jaillon O., Aury J.M., Brunet F., Petit J.L., Stange-Thomann N.,
RA Mauceli E., Bouneau L., Fischer C., Ozouf-Costaz C., Bernot A.,

```

RA Nicaud S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segurens B.,
 RA Dasilva C., Salancubet M., Levy M., Boudet N., Castellano S.,
 RA Anthouard V., Jubin C., Castelli V., Katinka M., Vacherie B.,
 RA Biemont C., Skalli Z., Catolico L., Poulain J., De Berardinis V.,
 RA Cruaud C., Duprat S., Brottier P., Coutanceau J.P., Gouzy J.,
 RA Parra G., Lardier G., Chapple C., McKernan K.J., McEwan P., Bosak S.,
 RA Kellis M., Volff J.N., Guigo R., Zody M.C., Mesirov J.,
 RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
 RA Laudet V., Schachter V., Quetier F., Saurin W., Scarpelli C.,
 RA Wincker P., Lander E.S., Weissbach J., Roest Crolius H.,
 RA "Genome duplication in the teleost fish Tetraodon nigroviridis reveals
 RT the early vertebrate proto-karyotype.";
 RL Nature 431:946-957(2004).
 RN (2)

NUCLEOTIDE SEQUENCE.

RG Genoscope; Whitehead Institute Centre for Genome Research;
 RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.

CC -!- CAUTION: The sequence shown here is derived from an

CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is

CC preliminary data.

DR EMBL; CAAB01014615; CAG00903.1; -!- Genomic DNA.

SQ SEQUENCE 1558 AA; 179215 MW; 2B3AF13FC3DFB39D CRC64;

Query Match 9.7%; Score 95.5; DB 2; Length 1558;

Best Local Similarity 25.1%; Pred. NO. 59;

Matches 42; Conservative 32; Mismatches 60; Indels 33; Gaps 7;

Qy 17 LVSHPVLTQGEAGRCADCEVCKEFLNRFYKSLIDRGVNFSLDTIEKELISFCLDTKG 76

Db 1097 LVS---LTEEEKIQNLVSPCEE-KEALQSSLSL-----NGEKEELQSLVSLCEEKKA 1147

Qy 77 KENRLCYVLGATKDAATKILSEVTRPMSVHPAMKICEKLLKLDSDICEKYEK---TLD 133

Db 1148 LQNRVY-----LSGDREKLRNHL--NFVGEQKKKLLKQLSSLSSEKEELQKD 1193

Qy 134 LASVDLRMRVAKLQILHSWG-----EECRACAEKTDYVNLQIE 173

Db 1194 LETLRQEKQLSAPRELPSNRGGQREAEQLQALQAEERSRRCSSLQE 1240

Search completed: December 13, 2005, 03:09:55

Job time : 233 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 13, 2005, 03:02:52 ; Search time 166 Seconds
(without alignments)
470.687 Million cell updates/sec

Title: US-10-648-361-2

Perfect score: 983

Sequence: 1 MWCASPVAVAFAGLLVSH.....VNLIQELAPKYAATHPKTEL 187

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

Published Applications AA Main:
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/US11_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	968	98.5	187	4	US-10-302-172-949
2	546	55.5	179	4	US-10-102-265-14
3	537	54.6	179	4	US-10-102-265-5
4	535	54.4	179	4	US-10-102-265-2
5	535	54.4	234	4	US-10-102-265-1
6	534.5	54.4	233	4	US-10-699-156-18
7	532.5	54.2	158	4	US-10-102-265-6
8	532.5	54.2	159	4	US-10-102-265-7
9	532.5	54.2	198	3	US-09-815-153-5
10	532.5	54.2	198	4	US-10-428-487-22
11	532	54.1	234	4	US-10-428-487-1
12	532	54.1	234	4	US-10-302-172-950
13	532	54.1	235	3	US-09-815-153-6
14	532	54.1	235	4	US-10-428-487-84
15	530.5	54.0	196	3	US-09-815-153-7
16	529.5	53.9	158	4	US-10-102-265-3
17	529.5	53.9	159	4	US-10-102-265-4
18	529	53.8	187	3	US-09-815-153-3
19	529	53.8	187	4	US-10-428-487-3
20	518	52.7	157	5	US-10-476-447-5
21	518	52.7	179	4	US-10-102-265-15
22	419	42.6	106	3	US-09-764-891-4527
23	393	40.0	173	6	US-11-097-143-41400
24	375.5	38.2	172	3	US-09-815-153-8
25	375.5	38.2	172	4	US-10-428-487-19
26	213	21.7	52	4	US-10-102-265-12
27	186	18.9	52	4	US-10-102-265-13

```

28      176      17.9      48      5      US-10-476-447-2      Sequence 2, Appli
29      143      14.5      42      5      US-10-476-447-3      Sequence 3, Appli
30     125.5     12.8      43      4      US-10-102-265-11     Sequence 11, Appl
31     112      11.4      39      4      US-10-106-698-6587     Sequence 6587, Ap
32     110.5     11.2      39      5      US-10-476-447-4      Sequence 4, Appli
33     101      10.3      507      4      US-10-109-670-12     Sequence 12, Appl
34      93       9.5      507      4      US-10-109-670-34     Sequence 34, Appl
35     90.5      9.2      331      4      US-10-464-261-4      Sequence 4, Appli
36     90.5      9.2     1027      4      US-10-654-416-8      Sequence 8, Appli
37     89.5      9.1      331      4      US-10-464-261-3      Sequence 3, Appli
38     89.5      9.1      775      4      US-10-125-770-4      Sequence 4, Appli
39     89.5      9.1     1029      4      US-10-654-416-6      Sequence 6, Appli
40     86.5      8.8      267      4      US-10-437-963-190696 Sequence 190696,
41     85.5      8.7      438      4      US-10-437-963-117135 Sequence 117135,
42     84.5      8.6      324      4      US-10-424-599-197017 Sequence 197017,
43     84.5      8.6     2621      4      US-10-437-963-122168 Sequence 122168,
44      84       8.5      683      5      US-10-515-477-18     Sequence 18, Appl
45      84       8.5      734      4      US-10-424-599-146660 Sequence 146660,

```

ALIGNMENTS

RESULT 1

```

US-10-302-172-949
; Sequence 949, Application US/10302172
; Publication No. US20040053250A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Dmanac, Radoje T.
; TITLE OF INVENTION: No. US20040053250A1el Arginine-rich Protein-like Nucleic Acids
; TITLE OF INVENTION: Polytypeptides
; FILE REFERENCE: 803 ICNCP
; CURRENT APPLICATION NUMBER: US/10/302,172
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/225,251
; PRIOR FILING DATE: 2002-08-20
; PRIOR APPLICATION NUMBER: PCT US02/05095
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: US 09/799,451
; PRIOR FILING DATE: 2001-03-05
; NUMBER OF SEQ ID NOS: 950
; SOFTWARE: pf_fl_genes Version 2.0
; SEQ ID NO 949
; LENGTH: 187
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-302-172-949

```

Query Match 98.5%; Score 968; DB 4; Length 187;

Best Local Similarity 98.9%; Pred. No. 2.1e-94;

Matches 185; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MWCASPVAVAFAGLLVSHPVLTQGOAGCPGADCEVCKEFLNRFYKSLIDRGVNSL 60

Db 1 MWCASPVAVAFAGLLVSHPVLTQGOAGCPGADCEVCKEFLNRFYKSLIDRGVNSL 60

Qy 61 DTIEKELISFCLDTKGKGNRLCYVYGATKDAATKILSEVTRPMSVHMPAMKICEKLLKD 120

Db 61 DTIEKELISFCLDTKGKGNRLCYVYGATKDAATKILSEVTRPMSVHMPAMKICEKLLKD 120

Qy 121 SQICELKYKTKLDIASVDLRKMRVAELKQILHSGECCRAAEKTDYVNLIOELAPKYAA 180

Db 121 SQICELKYKTKLDIASVDLRKMRVAELKQILHSGECCRAAEKTDYVNLIOELAPKYAA 180

Qy 181 THPKTEL 187

Db 181 THPKTEL 187

RESULT 2

US-10-102-265-14

```
/ Sequence 14, Application US/10102265
; Publication No. US20020182198A1
; GENERAL INFORMATION:
; APPLICANT: Commissiong, John W.
; TITLE OF INVENTION: Dopaminergic Neuronal Survival-Promoting
; TITLE OF INVENTION: Factors and Uses Thereof
; FILE REFERENCE: 50097/011002
; CURRENT APPLICATION NUMBER: US/10/102,265
; CURRENT FILING DATE: 2002-03-20
; PRIOR APPLICATION NUMBER: US 60/277,516
; PRIOR FILING DATE: 2001-03-20
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 179
; TYPE: PRT
; ORGANISM: Bos taurus
US-10-102-265-14

Query Match 55.5%; Score 546; DB 4; Length 179;
Best Local Similarity 56.6%; Pred. No. 1.5e-49;
Matches 107; Conservative 28; Mismatches 42; Indels 12; Gaps 3;

QY 1 MWCASPVAVVAFCAAGLLVSHPVLTQGEAGRGADCEVCCKEFLNRFYKSLIDRGVNFSL 60
DB 1 MWATHGLA-VALALSVPASRALRQ-----DCEVCISYLGFRFYQDLKDRDVTFS 50

QY 61 DTIEKELISFCLDTKGKCNRLCYLGGATKDAATKILSEVTRPMSPVHMPAMKICEKLLKLD 120
DB 51 ASIEKELIKFCREARGKCNRLCYIGATDDAATKIINEVSKPLSHHIPVEKICEKLLKLD 110

QY 121 SQICELKYKTDLASVDLKMVAELKQILHSGWGEACRAEKTQYVNLIOELAPKYA- 179
DB 111 SQICELKYKQIDLSTVDLKKRVKELKILDDWGETCKGCAEKSDYIRKINELMPKYAP 170

QY 180 -ATHPKTEL 187
DB 171 KAASRTDL 179

US-10-102-265-15

Query Match 54.4%; Score 535; DB 4; Length 179;
Best Local Similarity 56.6%; Pred. No. 2.3e-48;
Matches 107; Conservative 25; Mismatches 45; Indels 12; Gaps 4;

QY 1 MWCASPVAVVAFCAAGLLVSHPVLTQGEAGRGADCEVCCKEFLNRFYKSLIDRGVNFSL 60
DB 1 MWATQGLA-VALALSVPASRAL-----RPG-DCEVCISYLGFRFYQDLKDRDVTFS 50

QY 61 DTIEKELISFCLDTKGKCNRLCYLGGATKDAATKILSEVTRPMSPVHMPAMKICEKLLKLD 120
DB 51 ATIENELIKFCREARGKCNRLCYIGATDDAATKIINEVSKPLSHHIPVEKICEKLLKLD 110

QY 121 SQICELKYKTDLASVDLKMVAELKQILHSGWGEACRAEKTQYVNLIOELAPKYA- 179
DB 111 SQICELKYKQIDLSTVDLKKRVKELKILDDWGETCKGCAEKSDYIRKINELMPKYAP 170

QY 180 -ATHPKTEL 187
DB 171 KAASRTDL 179

US-10-102-265-16

Query Match 54.6%; Score 537; DB 4; Length 179;
Best Local Similarity 56.6%; Pred. No. 1.4e-48;
Matches 107; Conservative 27; Mismatches 43; Indels 12; Gaps 4;

QY 1 MWCASPVAVVAFCAAGLLVSHPVLTQGEAGRGADCEVCCKEFLNRFYKSLIDRGVNFSL 60
DB 1 MWATHGLA-VALALSVPASRAL-----RPG-DCEVCISYLGFRFYQDLKDRDVTFS 50

QY 61 DTIEKELISFCLDTKGKCNRLCYLGGATKDAATKILSEVTRPMSPVHMPAMKICEKLLKLD 120
DB 51 DTIEKELISFCLDTKGKCNRLCYLGGATKDAATKILSEVTRPMSPVHMPAMKICEKLLKLD 120
```


Db 3 RPQ-DCEVCISYLGREFYQDLKDRDVTFSPTATIEELIKFCREARGKENRLCYIYGATDDA 61
Qy 92 ATKILSEVTRPMSVHPMPAMKICEKLLKDSQICELKYKTKDLASVDLRMRVAELKQIL 151
Db 62 ATKIINEVSKPLAHHPVEKIKCKKDSQICELKYKQIDLSVTLKRLVKELKIL 121
Qy 152 HSWGECRACAEKTDVNLIOELAPKYA--ATHPKTEL 187
Db 122 DWGEMCKGCAEKSDYIRKINELMPKYAPKAASARTDL 159

RESULT 9
US-09-815-153-5
; Sequence 5, Application US/09815153
; Patent No. US20020132978A1
; GENERAL INFORMATION:
; APPLICANT: RASTELLI, LUCA K.
; APPLICANT: GERBER, HANS-PETER
; TITLE OF INVENTION: VEGF-MODULATED GENES AND METHODS EMPLOYING THEM
; FILE REFERENCE: 10716/34
; CURRENT APPLICATION NUMBER: US/09/815,153
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,201
; PRIOR FILING DATE: 2000-03-21
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 198
; TYPE: PRT
; ORGANISM: Murine sp.
US-09-815-153-5

Query Match 54.2%; Score 532.5; DB 3; Length 198;
Best Local Similarity 53.8%; Pred. No. 4.8e-48;
Matches 106; Conservative 29; Mismatches 37; Indels 25; Gaps 5;
Qy 15 GLLVSHPVLTQGE-----AGG-----RPGADCEVCCKEFLNRPYKSLID- 53
Db 3 GSFARHPAVVETAEEDVGVARAGGTLALSVLPSRALRPG-DCEVCISYLGREFYQDLVEG 61
Qy 54 -RGVNFSLDTIEKELISFCCLDTGKENRLCYILGATKDAATKILSEVTRPMSVHPMPAMKI 112
Db 62 FRDVTFSPTATIEELIKFCREARGKENRLCYIYGATDDAATKIINEVSKPLAHHPVEKI 121
Qy 113 CEKLLKDSQICELKYKTKDLASVDLRMRVAELKQILHSWGECRACAEKTDYVNLIO 172
Db 122 CEKLLKDSQICELKYKQIDLSVTLKRLVKELKILDDWGEMCKGCAEKSDYIRKIN 181
Qy 173 ELAPKYA--ATHPKTEL 187
Db 182 ELMPKYAPKAASARTDL 198

RESULT 10
US-10-428-487-22
; Sequence 22, Application US/10428487
; Publication No. US20040006780A1
; GENERAL INFORMATION:
; APPLICANT: RASTELLI, LUCA K.
; APPLICANT: GERBER, HANS-PETER
; TITLE OF INVENTION: VEGF-MODULATED GENES AND METHODS EMPLOYING THEM
; FILE REFERENCE: 0980080-0103
; CURRENT APPLICATION NUMBER: US/10/428,487
; CURRENT FILING DATE: 2003-05-02
; PRIOR APPLICATION NUMBER: 09/815,153
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,201
; PRIOR FILING DATE: 2000-03-22
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 22
; LENGTH: 198

; TYPE: PRT
; ORGANISM: Mus musculus
US-10-428-487-22
Query Match 54.2%; Score 532.5; DB 4; Length 198;
Best Local Similarity 53.8%; Pred. No. 4.8e-48;
Matches 106; Conservative 29; Mismatches 37; Indels 25; Gaps 5;
Qy 15 GLLVSHPVLTQGE-----AGG-----RPGADCEVCCKEFLNRPYKSLID- 53
Db 3 GSFARHPAVVETAEEDVGVARAGGTLALSVLPSRALRPG-DCEVCISYLGREFYQDLVEG 61
Qy 54 -RGVNFSLDTIEKELISFCCLDTGKENRLCYILGATKDAATKILSEVTRPMSVHPMPAMKI 112
Db 62 FRDVTFSPTATIEELIKFCREARGKENRLCYIYGATDDAATKIINEVSKPLAHHPVEKI 121
Qy 113 CEKLLKDSQICELKYKTKDLASVDLRMRVAELKQILHSWGECRACAEKTDYVNLIO 172
Db 122 CEKLLKDSQICELKYKQIDLSVTLKRLVKELKILDDWGEMCKGCAEKSDYIRKIN 181
Qy 173 ELAPKYA--ATHPKTEL 187
Db 182 ELMPKYAPKAASARTDL 198

RESULT 11
US-10-428-487-1
; Sequence 1, Application US/10428487
; Publication No. US20040006780A1
; GENERAL INFORMATION:
; APPLICANT: RASTELLI, LUCA K.
; APPLICANT: GERBER, HANS-PETER
; TITLE OF INVENTION: VEGF-MODULATED GENES AND METHODS EMPLOYING THEM
; FILE REFERENCE: 0980080-0103
; CURRENT APPLICATION NUMBER: US/10/428,487
; CURRENT FILING DATE: 2003-05-02
; PRIOR APPLICATION NUMBER: 09/815,153
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,201
; PRIOR FILING DATE: 2000-03-22
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 234
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-428-487-1

Query Match 54.1%; Score 532; DB 4; Length 234;
Best Local Similarity 56.1%; Pred. No. 6.8e-48;
Matches 106; Conservative 26; Mismatches 45; Indels 12; Gaps 3;
Qy 1 MWCASPVAVVAFCAGLLVSHPVLTQGEAGRRPGADCEVCCKEFLNRPYKSLIDRGVNFSL 60
Db 56 MMATQGLAVRVVALSVLPGRAL-----RPG-DCEVCISYLGREFYQDLKDRDVTFS 105
Qy 61 DTIEKELISFCCLDTGKENRLCYILGATKDAATKILSEVTRPMSVHPMPAMKICEKLLKLD 120
Db 106 ATIENELIKFCREARGKENRLCYIYGATDDAATKIINEVSKPLAHHPVEKICEKLLKLD 165
Qy 121 SQICELKYKTKDLASVDLRMRVAELKQILHSWGECRACAEKTDYVNLIOELAPKYA- 179
Db 166 SQICELKYKQIDLSVTLKRLVKELKILDDWGEMCKGCAEKSDYIRKINELMPKYAP 225
Qy 180 -ATHPKTEL 187
Db 226 KAASARTDL 234

RESULT 12
US-10-302-172-950
; Sequence 950, Application US/10302172
; Publication No. US20040053250A1

GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Dmanac, Radoje T.
; TITLE OF INVENTION: No. US20040053250A1el Arginine-rich Protein-like Nucleic Acids and
; FILE REFERENCE: 803 1CNCp
; CURRENT APPLICATION NUMBER: US/10/302,172
; CURRENT FILING DATE: 2002-11-21
; PRIOR FILING DATE: 2002-08-20
; PRIOR APPLICATION NUMBER: PCT US02/05095
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: US 09/799,451
; PRIOR FILING DATE: 2001-03-05
; NUMBER OF SEQ ID NOS: 950
; SOFTWARE: pc_fl_genes Version 2.0
; SEQ ID NO 950
; LENGTH: 234
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-302-172-950

Query Match 54.1%; Score 532; DB 4; Length 234;
Best Local Similarity 56.1%; Pred. No. 6.9e-48;
Matches 106; Conservative 26; Mismatches 45; Indels 12; Gaps 3;
Qy 1 MWCASPVAVAFACAGLLVSHPVLTQGEAGRGPGADCEVCKEFLNRFYKSLIDRGVNFSL 60
Db 56 MWATQGLAVRVALSVLPGRAL-----RPG-DCEVCISYLGRFYQDLKDRDVTFS 105
Qy 61 DTIEKELISFCLDTKGKRNLCYYLGATKDAATKILSEVTRPMSVHMPAMKICEKLLKLD 120
Db 106 ATIENELIKFCREARGKENRLCYIYGATDDAATKIINEVSKPLAHHIPVEKICEKLLKLD 165
Qy 121 SQICELKYKTLDLASVDRKMRVAELKQIILSHSGEGRACAEKTDYVNLIOELAPKYA- 179
Db 166 SQICELKYDKQIDLTSTVDLKKLRVKELKILDDNGETCKGCAEKSDYIRKINELMPKYAP 225
Qy 180 -ATHPKTEL 187
Db 226 KAASAPTDL 234

RESULT 13
US-09-815-153-6
; Sequence 6, Application US/09815153
; Patent No. US20020132978A1
; GENERAL INFORMATION:
; APPLICANT: RASTELLI, LUCA K.
; APPLICANT: GERBER, HANS-PETER
; TITLE OF INVENTION: VEGF-MODULATED GENES AND METHODS EMPLOYING THEM
; FILE REFERENCE: 10716/34
; CURRENT APPLICATION NUMBER: US/09/815,153
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,201
; PRIOR FILING DATE: 2000-03-21
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 235
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-815-153-6

Query Match 54.1%; Score 532; DB 3; Length 235;
Best Local Similarity 56.1%; Pred. No. 6.9e-48;
Matches 106; Conservative 26; Mismatches 45; Indels 12; Gaps 3;
Qy 1 MWCASPVAVAFACAGLLVSHPVLTQGEAGRGPGADCEVCKEFLNRFYKSLIDRGVNFSL 60
Db 57 MWATQGLAVRVALSVLPGRAL-----RPG-DCEVCISYLGRFYQDLKDRDVTFS 106

Qy 61 DTIEKELISFCLDTKGKRNLCYYLGATKDAATKILSEVTRPMSVHMPAMKICEKLLKLD 120
Db 107 ATIENELIKFCREARGKENRLCYIYGATDDAATKIINEVSKPLAHHIPVEKICEKLLKLD 166
Qy 121 SQICELKYKTLDLASVDRKMRVAELKQIILSHSGEGRACAEKTDYVNLIOELAPKYA- 179
Db 167 SQICELKYDKQIDLTSTVDLKKLRVKELKILDDNGETCKGCAEKSDYIRKINELMPKYAP 226
Qy 180 -ATHPKTEL 187
Db 227 KAASAPTDL 235
RESULT 14
US-10-428-487-84
; Sequence 84, Application US/10428487
; Publication No. US20040006780A1
; GENERAL INFORMATION:
; APPLICANT: RASTELLI, LUCA K.
; APPLICANT: GERBER, HANS-PETER
; TITLE OF INVENTION: VEGF-MODULATED GENES AND METHODS EMPLOYING THEM
; FILE REFERENCE: 09800080-0103
; CURRENT APPLICATION NUMBER: US/10/428,487
; CURRENT FILING DATE: 2003-05-02
; PRIOR APPLICATION NUMBER: 09/815,153
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,201
; PRIOR FILING DATE: 2000-03-22
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 84
; LENGTH: 235
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-428-487-84

Query Match 54.1%; Score 532; DB 4; Length 235;
Best Local Similarity 56.1%; Pred. No. 6.9e-48;
Matches 106; Conservative 26; Mismatches 45; Indels 12; Gaps 3;
Qy 1 MWCASPVAVAFACAGLLVSHPVLTQGEAGRGPGADCEVCKEFLNRFYKSLIDRGVNFSL 60
Db 57 MWATQGLAVRVALSVLPGRAL-----RPG-DCEVCISYLGRFYQDLKDRDVTFS 106
Qy 61 DTIEKELISFCLDTKGKRNLCYYLGATKDAATKILSEVTRPMSVHMPAMKICEKLLKLD 120
Db 107 ATIENELIKFCREARGKENRLCYIYGATDDAATKIINEVSKPLAHHIPVEKICEKLLKLD 166
Qy 121 SQICELKYKTLDLASVDRKMRVAELKQIILSHSGEGRACAEKTDYVNLIOELAPKYA- 179
Db 167 SQICELKYDKQIDLTSTVDLKKLRVKELKILDDNGETCKGCAEKSDYIRKINELMPKYAP 226
Qy 180 -ATHPKTEL 187
Db 227 KAASAPTDL 235
RESULT 15
US-09-815-153-7
; Sequence 7, Application US/09815153
; Patent No. US20020132978A1
; GENERAL INFORMATION:
; APPLICANT: RASTELLI, LUCA K.
; APPLICANT: GERBER, HANS-PETER
; TITLE OF INVENTION: VEGF-MODULATED GENES AND METHODS EMPLOYING THEM
; FILE REFERENCE: 10716/34
; CURRENT APPLICATION NUMBER: US/09/815,153
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,201
; PRIOR FILING DATE: 2000-03-21
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7

```

; LENGTH: 196
; TYPE: PRT
; ORGANISM: Murine sp.
4 US-09-815-153-7

Query Match      54.0%; Score 530.5; DB 3; Length 196;
Best Local Similarity 53.8%; Pred. No. 7.7e-48;
Matches 105; Conservative 27; Mismatches 40; Indels 23; Gaps 4;

QY 15 GLLVSHPVLTQGE-----AGG-----RPGADCEVCCKEFLNRFYKSLIDR 54
Db 3 GSFARHPAVVETAEDVGVARAGGTLALSVLPSRALRPG-DCEVCISYLGREFYQDLKDR 61

QY 55 GVNFSLDTIKELISFCLDTKGENRLCYVLGATKDAATKILSEVTRPMSVHMPAMKICE 114
Db 62 DVTFSPATIEELIKFCREARGKERNLCYYIGATDDAATKIINEVSKPLAHHIPVEKICE 121

QY 115 KLKLDQSQICELKYKTLDLASVDLRMRVAELKQILHSGWGECCACAEKTDYVNLIQEL 174
Db 122 KLKLDQSQICELKYDKQIDLSTVDLKKLRVKELKKILDDWGENCKGCAEKSDYIRKINEL 181

QY 175 APXYA--ATHPKTEL 187
Db 182 MPKYAPKAASARTDL 196

```

Search completed: December 13, 2005, 03:14:20
Job time : 167 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 13, 2005, 03:06:07 ; Search time 11 Seconds
(without alignments)
94.935 Million cell updates/sec

Title: US-10-648-361-2

Perfect score: 983

Sequence: 1 MWCASPVAVVAFCAAGLLVSH.....VNLIQELAPKYAATHPKTEL 187

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 32527 seqs, 5584426 residues

Total number of hits satisfying chosen parameters: 32527

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

Published Applications_AA New:*

1: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	538	54.7	182	6	US-10-821-234-1420
2	82.5	8.4	355	7	US-11-108-528-78
3	82.5	8.4	365	7	US-11-108-528-76
4	79.5	8.1	394	6	US-10-821-234-1187
5	79	8.0	524	6	US-10-821-234-1631
6	74	7.5	333	7	US-11-004-789-2
7	72.5	7.4	643	6	US-10-510-386-8
8	72	7.3	296	6	US-10-131-826A-26
9	72	7.3	733	7	US-11-012-762-68
10	71.5	7.3	1062	7	US-11-137-465-43
11	70.5	7.2	977	7	US-10-993-274-39
12	69.5	7.1	636	6	US-10-467-657-1856
13	69.5	7.1	718	7	US-11-074-176-306
14	69.5	7.1	723	7	US-11-074-176-18
15	67.5	6.9	319	7	US-11-074-176-74
16	67	6.8	250	6	US-10-821-234-1659
17	67	6.8	729	7	US-11-099-691-3
18	66.5	6.8	317	6	US-10-467-657-3320
19	66	6.7	1451	7	US-11-046-346-1
20	65.5	6.7	179	6	US-10-485-517-158
21	65.5	6.7	3002	6	US-10-821-234-916
22	65	6.6	480	6	US-10-878-556A-198
23	64.5	6.6	752	6	US-10-793-626-1138
24	64	6.5	353	7	US-11-012-762-50
25	64	6.5	777	6	US-10-467-657-2474

Sequence 27, Appl
Sequence 24, Appl
Sequence 211, Appl
Sequence 11, Appl
Sequence 179, Appl
Sequence 366, Appl
Sequence 57, Appl
Sequence 1396, Ap
Sequence 1188, Ap
Sequence 428, Ap
Sequence 216, Ap
Sequence 7912, Ap
Sequence 25, Appl
Sequence 122, Ap
Sequence 195, Ap
Sequence 196, Ap
Sequence 328, Ap
Sequence 76, Appl
Sequence 1358, Ap
Sequence 15, Appl

63.5 6.5 829 6 US-10-512-109-27
63.5 6.5 1531 7 US-11-087-227-24
63.5 6.5 1531 7 US-11-186-284-211
63 6.4 295 7 US-11-067-121-11
63 6.4 354 6 US-10-485-517-179
63 6.4 354 6 US-10-485-517-366
63 6.4 883 6 US-10-770-726-57
62.5 6.4 301 6 US-10-793-626-1396
62.5 6.4 930 6 US-10-821-234-1188
62 6.3 317 7 US-11-000-463-428
62 6.3 321 6 US-10-467-657-216
62 6.3 321 6 US-10-467-657-7912
62 6.3 407 7 US-11-051-267-25
62 6.3 476 7 US-11-074-176-122
61.5 6.3 187 7 US-11-186-284-195
61.5 6.3 346 6 US-10-467-657-2676
61.5 6.3 379 6 US-10-131-826A-328
61.5 6.3 450 6 US-10-763-712A-76
61.5 6.3 583 6 US-10-793-626-1358
61 6.2 318 7 US-11-109-156-15

ALIGNMENTS

RESULT 1
US-10-821-234-1420
; Sequence 1420, Application US/10821234
; Publication No. US20050255114A1
; GENERAL INFORMATION:
; APPLICANT: Labat, Ivan
; APPLICANT: Stache-Crain, Birgit
; APPLICANT: Andarmani, Susan
; APPLICANT: Tang, Y. Tom
; TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia
; FILE REFERENCE: 821A
; CURRENT APPLICATION NUMBER: US/10/821.234
; CURRENT FILING DATE: 2004-04-07
; PRIOR FILING DATE: 2003-04-07
; NUMBER OF SEQ ID NOS: 1704
; SOFTWARE: PT_SEQ_genes Version 1.0
; SEQ ID NO 1420
; LENGTH: 182
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-821-234-1420

Query Match 54.7% Score 538; DB 6; Length 182;
Best Local Similarity 56.6%; Pred No. 1e-46;
Matches 107; Conservative 26; Mismatches 44; Indels 12; Gaps 4;
Qy 1 MWCASPVAVVAFCAAGLLVSHPVLTQGOAGGRCPCGADCEVCCKEFLNRFYKSLIDRGVNFSL 60
Db 4 MWATQGLA-VALALSVLPGRAL-----RPG-DCEVCISYLGRFYQDLKDRDVTFSF 53
Qy 61 DTEKELISCLDTKGNELCYLGGATKDAATKILSEVTRPMVHMPAMKICSLKLLD 120
Db 54 ATIENELIKFCREARGENELCYIYGATDAATKINEVSKPLAHIPVEKICSLKKKO 113
Qy 121 SQICELKYETLDIASVDLKRVAELKQILHSGEFCRACAEKTDYVNLIQELAPKYA- 179
Db 114 SQICELKYQIDLVTLKRLVKELKLLDDWGETCKGCAEKSDYIRKINELMPKYAP 173
Qy 180 -ATHPKTEL 187
Db 174 KAASARTDL 182

RESULT 2

US-11-108-528-78
; Sequence 78, Application US/11108528
; Publication No. US20050261189A1

```
/ GENERAL INFORMATION:
/ APPLICANT: Larsen, Glenn
/ APPLICANT: Marvin, Martha
/ APPLICANT: Li, Dean Y.
/ APPLICANT: Wang, Elizabeth
/ APPLICANT: Chen, C. M. Amy
/ APPLICANT: Shamah, Steven M.
/ TITLE OF INVENTION: METHODS OF PROMOTING CARDIAC CELL
/ FILE REFERENCE: HYDR-P01-041
/ CURRENT APPLICATION NUMBER: US/11/108,528
/ PRIOR FILING DATE: 2005-04-18
/ PRIOR APPLICATION NUMBER: US 60/563,137
/ PRIOR FILING DATE: 2004-04-16
/ PRIOR APPLICATION NUMBER: US 60/598,368
/ PRIOR FILING DATE: 2004-08-02
/ NUMBER OF SEQ ID NOS: 86
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 78
/ LENGTH: 355
/ TYPE: PRT
/ ORGANISM: Homo sapiens
/
US-11-108-528-78

Query Match      8.4%; Score 82.5; DB 7; Length 355;
Best Local Similarity 23.0%; Pred. No. 0.3;
Matches 50; Conservative 28; Mismatches 58; Indels 71; Gaps 10;

QY 8 AVAFACAGLLVSHPVITQGEAGRPADCEV-----CKEFLNRYKSL 51
Db 115 AVMA--AGLVHS---VTRSCAGNMTECSDTTLQNGSASGHWGGSDD----- 161
QY 52 IDRGVNFSLDTIEKELISFCL-DTKGENRLCYLGG-----ATKDAATKILSEVTRPSV 105
Db 162 VOYGMWFS-----RKFLDPFGNTGKENVLLAMNLHNEAGRQAVAKLMSVDCRCHGV 216
QY 106 H-MPAMKICBKLKLDQSICEL---KYEKTLDLASVDLRKMRVAELKQ----- 149
Db 217 SGSCAVKTCWKTMSSFEKIGHLLKDKYENSIQISDKTKRMRRREKDKRKIPHKDDLLY 276
QY 150 -----ILHSWGEECRACAEKTDYVNL 171
Db 277 VNKSPNYCYVEDKKLGIPGTQGRECNRTSEGADGCNLL 313

RESULT 3
US-11-108-528-76
/ Sequence 76, Application US/11/108528
/ Publication No. US20050261189A1
/ GENERAL INFORMATION:
/ APPLICANT: Larsen, Glenn
/ APPLICANT: Marvin, Martha
/ APPLICANT: Li, Dean Y.
/ APPLICANT: Wang, Elizabeth
/ APPLICANT: Chen, C. M. Amy
/ APPLICANT: Shamah, Steven M.
/ TITLE OF INVENTION: METHODS OF PROMOTING CARDIAC CELL
/ FILE REFERENCE: HYDR-P01-041
/ CURRENT APPLICATION NUMBER: US/11/108,528
/ PRIOR FILING DATE: 2005-04-18
/ PRIOR APPLICATION NUMBER: US 60/563,137
/ PRIOR FILING DATE: 2004-04-16
/ PRIOR APPLICATION NUMBER: US 60/598,368
/ PRIOR FILING DATE: 2004-08-02
/ NUMBER OF SEQ ID NOS: 86
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 76
/ LENGTH: 365
/ TYPE: PRT
/ ORGANISM: Homo sapiens
/
US-11-108-528-76

Query Match      8.4%; Score 82.5; DB 7; Length 365;
Best Local Similarity 23.0%; Pred. No. 0.31;
Matches 50; Conservative 28; Mismatches 58; Indels 71; Gaps 10;

QY 8 AVAFACAGLLVSHPVITQGEAGRPADCEV-----CKEFLNRYKSL 51
Db 125 AVMA--AGLVHS---VTRSCAGNMTECSDTTLQNGSASGHWGGSDD----- 171
QY 52 IDRGVNFSLDTIEKELISFCL-DTKGENRLCYLGG-----ATKDAATKILSEVTRPSV 105
Db 172 VOYGMWFS-----RKFLDPFGNTGKENVLLAMNLHNEAGRQAVAKLMSVDCRCHGV 226
QY 106 H-MPAMKICBKLKLDQSICEL---KYEKTLDLASVDLRKMRVAELKQ----- 149
Db 227 SGSCAVKTCWKTMSSFEKIGHLLKDKYENSIQISDKTKRMRRREKDKRKIPHKDDLLY 286
QY 150 -----ILHSWGEECRACAEKTDYVNL 171
Db 287 VNKSPNYCYVEDKKLGIPGTQGRECNRTSEGADGCNLL 323

RESULT 4
US-10-821-234-1187
/ Sequence 1187, Application US/10821234
/ Publication No. US20050255114A1
/ GENERAL INFORMATION:
/ APPLICANT: Labat, Ivan
/ APPLICANT: Stache-Crain, Birgit
/ APPLICANT: Andarmani, Susan
/ APPLICANT: Tang, Y. Tom
/ TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia
/ FILE REFERENCE: 821A
/ CURRENT APPLICATION NUMBER: US/10/821,234
/ CURRENT FILING DATE: 2004-04-07
/ PRIOR APPLICATION NUMBER: US 60/462,047
/ PRIOR FILING DATE: 2003-04-07
/ NUMBER OF SEQ ID NOS: 1704
/ SOFTWARE: pt_seq_genes Version 1.0
/ SEQ ID NO 1187
/ LENGTH: 394
/ TYPE: PRT
/ ORGANISM: Homo sapiens
/
US-10-821-234-1187

Query Match      8.1%; Score 79.5; DB 6; Length 394;
Best Local Similarity 25.0%; Pred. No. 0.67;
Matches 44; Conservative 21; Mismatches 42; Indels 69; Gaps 10;

QY 67 LISFCLDTGKGNRLCYLGG-----ATKDA--ATKILSEVTRPSVHM-----PAMX 111
Db 118 LLGECMIRHKE-----LGESNFGDALLDAGESMKRLAEVKDSLIDIEVKQNFIDPLQN 171
QY 112 ICEK-----LKKLDSQICELKYBKTLDLASVDLRKMRVAELKQILHSWGEECRACAE 163
Db 172 LCEKDLKEIQHHLKLEGRRLDFYKKRQ-----GKIPDEELQALEKF-BESKEVAE 224
QY 164 -----KTDY-----VNLIQELAPKY-----AATHPKTE 186
Db 225 TSMHNLLETDTIEQVSQLSALVDQAQLDYHQAQVQILDLEAEKLRMRREASSRPRRE 280

RESULT 5
US-10-821-234-1631
/ Sequence 1631, Application US/10821234
/ Publication No. US20050255114A1
/ GENERAL INFORMATION:
/ APPLICANT: Labat, Ivan
/ APPLICANT: Stache-Crain, Birgit
/ APPLICANT: Andarmani, Susan
/ APPLICANT: Tang, Y. Tom
/ TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia
/ FILE REFERENCE: 821A
/ CURRENT APPLICATION NUMBER: US/10/821,234
```

```

; CURRENT FILING DATE: 2004-04-07
; PRIOR APPLICATION NUMBER: US 60/462,047
; PRIOR FILING DATE: 2003-04-07
; NUMBER OF SEQ ID NOS: 1704
; SOFTWARE: pt SEQ_genes Version 1.0
; SEQ ID NO 1631
; LENGTH: 524
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-821-234-1631

```

```

Query Match      8.0%; Score 79; DB 6; Length 524;
Best Local Similarity 20.4%; Pred. No. 1.1;
Matches 40; Conservative 35; Mismatches 65; Indels 56; Gaps 11;

Qy 1 MWC-----AGPVAVAFAGLLVSHPVLTQOEAGGRPGADCEVCCKEFLNRYFKSLIDRG 55
      |||      |||      |||      |||      |||      |||      |||      |||
Db 31 VMQNVKTSADCGAVKHCLQTVWKPPIV-----KSLPCDICKDVVTAAGDMLKD-- 79

Qy 56 VNFSLDTIRKELISFCIDT-----KGKENRLC-----YYLGATKDAATKILSEVTRPMSV 105
      |||      |||      |||      |||      |||      |||      |||      |||
Db 80 -----NATEBEILVYLEKTDWLPKPNMSASCKEIVDSYLPVILDI---IKGEMSRPGEV 131

Qy 106 HMPAMKICEKLKLDSDOICELKYEKTLTD---LASVDLRKM---RVAELKQILHSW----- 154
      |||      |||      |||      |||      |||      |||      |||      |||
Db 132 -CSALNLCESLQK---HLAEINHQKUESNKIPELDMTEVVAPPFMANIPILLVPQDGPRS 187

Qy 155 -----GEECRACAE 163
      |||      |||      |||
Db 188 KPQPKNGDVCODCIO 203
      |||      |||      |||

```

```

RESULT 6
US-11-004-789-2
; Sequence 2, Application US/11004789
; Publication No. US20050250941
; GENERAL INFORMATION:
; APPLICANT: Steig, Bernd
; APPLICANT: Yang, Maria X. H.
; TITLE OF INVENTION: MITOGEN-ACTIVA
; TITLE OF INVENTION: MEK6 AND METH
; FILE REFERENCE: 860098.403C1
; CURRENT APPLICATION NUMBER: US/11/
; CURRENT FILING DATE: 2004-12-03
; PRIOR APPLICATION NUMBER: US/09/59
; PRIOR FILING DATE: 2000-06-13
; PRIOR APPLICATION NUMBER: US 08/57
; PRIOR FILING DATE: 1995-12-20
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Vers
; SEQ ID NO 2
; LENGTH: 333
; TYPE: PRT
; ORGANISM: homo sapien
US-11-004-789-2

```

	Query Match	7.5%;	Score 74;	DB 7;	Length 333;
	Best Local Similarity	17.1%;	Pred. No. 1.9;		
	Matches	30;	Conservative 38;	Mismatches 59;	Indels 48; Gaps 5;
Qy	36	DCEVCKEF	---	LNRRFYKSLIDRGVNFSLDTIEKELISFCFLDTYKGRNRLCYVLGATKDA	91
Db	123	DWICHELMDTS	LQKFPYQYVIDGQTIPEDILGKIADVIVKALEHLHSLK	---	172
Qy	92	ATKLSEVTRPMSPVHMPAM	---	KICE	---
Db	173	--SVIHRDVKPSNVINALQVQKMD	FGISGLVDSVAKTIDAGCKPYMAPERINPELNQ	---	230
Qy	126	LKTEKTLDLASVDLRKMRVAELKQIILHSWGEECRACAETDYNVLQELAPKYAA	180		
Db	231	KGYSVKSDLSWLGITMIELAFPPDYDSWGTPTFQOLKQ	---	VVEEPSQPLPA	279

```

RESULT 7
US-10-510-386-8
; Sequence 8, Application US/10510386
; Publication No. US20050244922A1
; GENERAL INFORMATION:
; APPLICANT: Andersen, Jens Tonne
; APPLICANT: Clausen, Ib Groth
; APPLICANT: Jorgensen, Steen Troels
; APPLICANT: Olsen, Peter Bjarke
; APPLICANT: Rasmussen, Michael Dolberg
; TITLE OF INVENTION: Improved Bacillus Host Cell
; FILE REFERENCE: 10294.204-US
; CURRENT APPLICATION NUMBER: US/10/510,386
; CURRENT FILING DATE: 2004-10-04
; NUMBER OF SEQ ID NOS: 248
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 8
; LENGTH: 643
; TYPE: PRT
; ORGANISM: Bacillus licheniformis
US-10-510-386-8

Query Match          7.4%; Score 72.5; DB 6; Length 643;
Best Local Similarity 23.2%; Pred. No. 6.2;
Matches 32; Conservative 28; Mismatches 59; Indels 19; Gaps 5;

Qy      38  EVCKEFLNRFYKSLIDRGV--NFSLDTIEKELISFCLDTGKENRLCYLYLGATKDAA TKI 95
Db      381  EQTDQSVERTLQKIDRGIIIDHTGSPLOKK-IKYALKTVGSSKHKYKKOMARDSATIL 439
Qy      96  LSHWGEECRACAEKTDYV 168
Db      440  KNDKSTLTF-----KPKGDTVLILAPYEEQTAAIAKTSIKRKNIKVYRFA 488
Qy      151  LHSWGEECRACAEKTDYV 168
Db      489  EKTFFEEIOKKIDEADV 506

```

RESULT 8
 US-131-826A-26
 ; Sequence 26, Application US/10131826A
 ; Publication No. US20050245730A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Baker, Kevin P.
 ; APPLICANT: Beresini, Maureen
 ; APPLICANT: DeForge, Laura
 ; APPLICANT: Desnoyers, Luc
 ; APPLICANT: Filvaroff, Ellen
 ; APPLICANT: Gao, Wei-Qiang
 ; APPLICANT: Gerritsen, Mary E.
 ; APPLICANT: Goddard, Audrey
 ; APPLICANT: Godowski, Paul J.
 ; APPLICANT: Gurney, Austin L.
 ; APPLICANT: Sherwood, Steven
 ; APPLICANT: Smith, Victoria
 ; APPLICANT: Stewart, Timothy A.
 ; APPLICANT: Tumas, Daniel
 ; APPLICANT: Watanabe, Colin K
 ; APPLICANT: Wood, William
 ; APPLICANT: Zhang, Zemin
 ; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC ACIDS ENCODING THE SAME
 ; FILE REFERENCE: P3330R1C128
 ; CURRENT APPLICATION NUMBER: US/10/131,826A
 ; CURRENT FILING DATE: 2002-04-24
 ; PRIOR APPLICATION NUMBER: 60/049911
 ; PRIOR FILING DATE: 1997-06-18
 ; PRIOR APPLICATION NUMBER: 60/056974
 ; PRIOR FILING DATE: 1997-08-26
 ; PRIOR APPLICATION NUMBER: 60/059113
 ; PRIOR FILING DATE: 1997-09-17
 ; PRIOR APPLICATION NUMBER: 60/059115

```
/ PRIOR FILING DATE: 1997-09-17
/ PRIOR APPLICATION NUMBER: 60/059117
/ PRIOR FILING DATE: 1997-09-17
/ PRIOR APPLICATION NUMBER: 60/059122
/ PRIOR FILING DATE: 1997-09-17
/ PRIOR APPLICATION NUMBER: 60/059184
/ PRIOR FILING DATE: 1997-09-17
/ PRIOR APPLICATION NUMBER: 60/059263
/ PRIOR FILING DATE: 1997-09-18
/ PRIOR APPLICATION NUMBER: 60/059352
/ PRIOR FILING DATE: 1997-09-19
/ PRIOR APPLICATION NUMBER: 60/059588
/ PRIOR FILING DATE: 1997-09-19
/ Remaining Prior Application data removed - See File Wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 550
/ SEQ ID NO 26
/ LENGTH: 296
/ TYPE: PRT
/ ORGANISM: Homo Sapien
US-10-131-826A-26

Query Match      7.3%; Score 72; DB 6; Length 296;
Best Local Similarity 26.6%; Pred. No. 2.6;
Matches 41; Conservative 22; Mismatches 45; Indels 46; Gaps 9;

QY 57 NPSLDTIEKELISFCLDTKGKRNLCYVLGATKDAATKI-----LSEVTRPMS-- 104
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 98 NOTEDSLRKLIALQED---KHN-----YETAKESLRRVLQEKIEVVRKLSVEVSLSNT 150
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 105 ----VHMPAM-----KICEKLKLDQSICELKYEKTLDLASVDLRMRVAELKQILHSW 154
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 151 EDCETHLKNVNTQELRELANKYNGAVNEIK-----DLSD----KLKVAEGKQ----- 196
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 155 GESEC--RACAEKTDYVNLQELAPKYAATHPKTE 186
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 197 -EBIQKGQAEKKELQHKIDEMEKEQELQAKTE 229
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

RESULT 9
US-11-012-762-68
/ Sequence 68, Application US/11012762
/ Publication No. US20050244815A1
/ GENERAL INFORMATION:
/ APPLICANT: Georgia State University Research Foundation, Inc.
/ TITLE OF INVENTION: Compositions and Methods for Viral Resistance Genes
/ FILE REFERENCE: GSUI.PCT
/ CURRENT APPLICATION NUMBER: US/11/012,762
/ CURRENT FILING DATE: 2004-12-15
/ PRIOR APPLICATION NUMBER: PCT/US03/19300
/ PRIOR FILING DATE: 2003-06-19
/ PRIOR APPLICATION NUMBER: US 60/390,046
/ PRIOR FILING DATE: 2002-06-19
/ NUMBER OF SEQ ID NOS: 133
/ SOFTWARE: Patentin version 3.2
/ SEQ ID NO 68
/ LENGTH: 733
/ TYPE: PRT
/ ORGANISM: Rattus norvegicus
US-11-012-762-68

Query Match      7.3%; Score 72; DB 7; Length 733;
Best Local Similarity 19.0%; Pred. No. 8.2;
Matches 28; Conservative 35; Mismatches 60; Indels 24; Gaps 6;

QY 38 EVCKEFLNRRYKSLIDRGVNFSLDT-IEKELISFCLDTKGKRNLCYVLGATKDAATKIL 96
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 471 EACQKAFN-----LGVKFDISKWMAPRVLSFTLESKSLKQSVFDPVLPAYDALGQLR 522
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 97 SEVTRPMSVHPAMKIC---EKLLKLDQSICELKYEKTLDLASVDLRKMRVAELKQILHS 153
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 523 SDVTSRKAYKUIELVYASQSLKGGFSCFTLQKDF-----IETPTKLKGLIRLIK 578
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 154 WGESEC-RACAEKTDYVNLQELAPKYA 179
```

```
Db 579 WYQCEKCKMKPKA-----SLPPKYA 598
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

RESULT 10
US-11-137-465-43
/ Sequence 43, Application US/11137465
/ Publication No. US20050255558A1
/ GENERAL INFORMATION:
/ APPLICANT: Agarwal, Pankaj
/ APPLICANT: Murdoch, Paul R.
/ APPLICANT: Rizvi, Safia, K.
/ APPLICANT: Smith, Randall, F.
/ APPLICANT: Xiang, Zhaoying
/ APPLICANT: Kabnick, Karen
/ TITLE OF INVENTION: NOVEL COMPOUNDS
/ FILE REFERENCE: GP50018
/ CURRENT APPLICATION NUMBER: US/11/137,465
/ CURRENT FILING DATE: 2005-05-25
/ PRIOR APPLICATION NUMBER: US/10/239,663
/ PRIOR FILING DATE: 2002-09-24
/ PRIOR APPLICATION NUMBER: PCT/US01/09226
/ PRIOR FILING DATE: 2001-03-22
/ PRIOR APPLICATION NUMBER: 60/192,158
/ PRIOR FILING DATE: 2000-03-24
/ PRIOR APPLICATION NUMBER: 60/192,668
/ PRIOR FILING DATE: 2000-03-27
/ PRIOR APPLICATION NUMBER: 60/200,166
/ PRIOR FILING DATE: 2000-04-27
/ NUMBER OF SEQ ID NOS: 66
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 43
/ LENGTH: 1062
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-11-137-465-43

Query Match      7.3%; Score 71.5; DB 7; Length 1062;
Best Local Similarity 18.8%; Pred. No. 15;
Matches 25; Conservative 30; Mismatches 59; Indels 19; Gaps 4;

QY 55 GVNFSLDTIEKELISFCLDTKGKRNLC-----CYILGATKDAATKILSEVTRPMSVH 106
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 574 GCMSPD-IQELLRCDISCKGHSIVTDLQELLGLCLYESQEELVKEVMAQF-KKISLH 631
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 107 MPA-----MKICEKLKLDQSICELKYEKTLDLASVDLRKMRVAELKQILHSWGE 157
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 632 LNAVDVVPSSFCVKYKHCRLQKMSLQVIKENLPENVTASESDARVERSQDDQHMLPFTDL 691
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 158 CRACAEKTDYVNL 170
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 692 CSIFGSKDLMGL 704
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

RESULT 11
US-11-093-274-39
/ Sequence 39, Application US/11093274
/ Publication No. US20050266008A1
/ GENERAL INFORMATION:
/ APPLICANT: Graziano, Robert
/ APPLICANT: Cardarelli, Josephine M.
/ APPLICANT: Kempe, Thomas
/ APPLICANT: Cutter, Beth
/ APPLICANT: Srinivasan, Mohan
/ TITLE OF INVENTION: IRTA-5 ANTIBODIES AND THEIR USES
/ FILE REFERENCE: 04280/1201101-US1
/ CURRENT APPLICATION NUMBER: US/11/093,274
/ CURRENT FILING DATE: 2005-03-28
/ PRIOR APPLICATION NUMBER: 60/557,741
/ PRIOR FILING DATE: 2004-03-29
/ NUMBER OF SEQ ID NOS: 41
/ SOFTWARE: Patentin version 3.2
/ SEQ ID NO 39
```



```
; Publication No. US20050250135A1
; GENERAL INFORMATION:
; APPLICANT: Kleenhammer, Todd R.
; APPLICANT: Russell, William M.
; APPLICANT: Altermann, Eric
; APPLICANT: McAuliffe, Olivia
; APPLICANT: Peril, Andrea Azcarate
; TITLE OF INVENTION: Nucleic Acid Sequences Encoding
; TITLE OF INVENTION: Stress-Related Proteins and Uses Therefore
; FILE REFERENCE: 5051-694
; CURRENT APPLICATION NUMBER: US/11/074,176
; CURRENT FILING DATE: 2005-03-07
; PRIOR APPLICATION NUMBER: 60/551,161
; PRIOR FILING DATE: 2004-03-08
; NUMBER OF SEQ ID NOS: 381
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 74
; TYPE: PRT
; LENGTH: 319
; ORGANISM: Lactobacillus acidophilus
US-11-074-176-74
```

```
Query Match      6.9%; Score 67.5; DB 7; Length 319;
Best Local Similarity 26.7%; Pred. No. 7.9;
Matches 28; Conservative 12; Mismatches 52; Indels 13; Gaps 3;

QY      8 AVVAFGAG----LLVSHPVLTQGEAGGPGADCEVCCKEFLNRFYKSLID-RGVNFSLDT 62
Db      148 AIAAFKKGEVSRILTRPAVEAGESLGLFPGDLKEKVDPYLRPIYDSLYAILGTNTDRL 207

QY      63 IEKELISFCLDTGKENRLCYILGATKDAATKILSEVTRPMSVHM 107
Db      208 MERGVIEA-----PLAYMEGRTLDDAFVILDEAQNTTDAQM 244
```

```
Search completed: December 13, 2005, 03:14:38
Job time : 12 secs
```


GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 13, 2005, 02:57:42 ; Search time 37 Seconds
(without alignments)
486.285 Million cell updates/sec

Title: US-10-648-361-2
Perfect score: 983
Sequence: 1 MWCASPVAVAFCAAGLLVSH.....VNLQELAPKYAATHPKTEL 187

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_80:*
1: PIR1.*
2: PIR2.*
3: PIR3.*
4: PIR4.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	ID	Description
1	532	54.1	234	2 S27956	arginine-rich prot
2	95	9.7	498	1 S48058	cytochrome P450 Cy
3	89.5	9.1	588	1 BVEUC	excinuclease ABC,
4	89.5	9.1	588	2 D85808	excinuclease ABC,
5	89.5	9.1	588	2 C90960	excinuclease ABC s
6	89.5	9.1	775	2 I49237	A20 protein - mous
7	89	9.1	563	2 S13786	DNA-directed DNA p
8	88.5	9.0	828	2 T06133	hypothetical prote
9	88	9.0	39	2 S69268	arginine-rich prot
10	87.5	8.9	276	2 H81707	inclusion membrane
11	87.5	8.9	3587	2 I40486	surfactin syntheta
12	87	8.9	481	2 T10470	transcription init
13	87	8.9	583	2 T09157	phosphoglucutase
14	83	8.4	463	2 S26670	retinoic acid rece
15	83	8.4	463	2 B1727	retinoic acid rece
16	82	8.3	442	2 A38592	retinoic acid rece
17	82	8.3	470	2 D41977	retinoid receptor
18	82	8.3	475	2 B41977	retinoic acid rece
19	81.5	8.3	389	2 C49776	hypothetical prote
20	81.5	8.3	389	2 A71091	hypothetical prote
21	81.5	8.3	416	2 T25101	hypothetical prote
22	81.5	8.3	437	2 B70353	conserved hypotet
23	81.5	8.3	610	2 AG0227	excinuclease ABC c
24	81	8.2	328	2 H75073	hypothetical prote
25	81	8.2	366	2 E71920	hypothetical prote
26	81	8.2	455	2 A36471	transcription fact
27	80	8.1	435	2 T22332	hypothetical prote
28	80	8.1	719	2 S51739	transcription repr
29	80	8.1	880	2 F75103	conserved hypotet

30	80	8.1	1081	2 S15040	pleiotropic drug r
31	79.5	8.1	251	2 T04866	hypothetical prote
32	79.5	8.1	717	2 H72208	conserved hypotet
33	79.5	8.1	1127	2 T21635	hypothetical prote
34	79.5	8.1	1165	2 T21636	hypothetical prote
35	79.5	8.1	1750	2 H64403	ribonucleoside-tri
36	79	8.0	296	2 T41128	hypothetical prote
37	79	8.0	332	2 E70384	biotin synthase (E
38	79	8.0	465	2 S52074	hepatocyte nuclear
39	79	8.0	527	1 SAHUP	saposin precursor
40	79	8.0	715	2 JC2222	major surface glyco
41	79	8.0	726	2 T44825	hypothetical prote
42	79	8.0	863	2 H69527	valyl-tRNA synthet
43	79	8.0	879	2 C71083	conserved hypotet
44	79	8.0	1418	2 S64918	hypothetical prote
45	78.5	8.0	307	2 D84536	hypothetical prote

ALIGNMENTS

RESULT 1

S27956

arginine-rich protein - human

C:Species: Homo sapiens (man)

C>Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 09-Jul-2004

C:Accession: S27956

R:Golembieski, W.; Shridhar, V.; Miller, O.J.; Smith, D.I.

submitted to the EMBL Data Library, June 1992

A:Description: Identification of a new arginine-rich gene from a cosmid containing clus

A:Reference number: S27956

A:Accession: S27956

A:Molecule type: DNA

A:Residues: 1-234 <GOL>

A:Cross-references: UNIPROT:Q86U67; UNIPARC:UPI000015C44A; EMBL:M83751; NID:g1789930; PI

Query Match

Best Local Similarity 54.1%; Score 532; DB 2; Length 234;

Matches 106; Conservative 26; Mismatches 45; Indels 12; Gaps 3;

Qy 1 MWCASPVAVAFCAAGLLVSHPVLTQGOEAGRGPGADCEVCCKEFLNRFYKSLIDRGVNFSL 60

Db 56 MWATQGLAVRALVSLVPGSRAL-----RPG-DCEVCISILGRFVQDLKDRDVTSP 105

Qy 61 DTEKELISCLDTKGKRNLCYLGATKDAATKILSEVTRPMSVHMPAMKICSLKKLD 120

Db 106 ATTENELIKFCREARGKENRLCYIGATDDAATKIIINEVSKPLAHIPVEXICSLKKKD 165

Qy 121 SQICELKYEKTLDLASVDLRKMRVAELKQILHSWGECRACAEKTDYVNLQELAPKYA- 179

Db 166 SQICELKYQIDLVTLVKLRVLEKLDLDDWGETCKGCAEKSDYIRKINELMPKYAP 225

Qy 180 -ATHPKTEL 187

Db 226 KAASAPTDL 234

RESULT 2

S48058

cytochrome P450 CYP6B1 - black swallowtail

N:Contains: oxidoreductase (EC 1.-.-.-)

C:Species: Papilio polyxenes (black swallowtail)

C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004

C:Accession: S48058; A46367

R:Prapaipong, H.; Berenbaum, M.R.; Schuler, M.A.

Nucleic Acids Res. 22, 3210-3217, 1994

A:Title: Transcriptional regulation of the Papilio polyxenes CYP6B1 gene.

A:Reference number: S48058; MUID:94344788; PMID:8065937

A:Accession: S48058

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-498 <PRA>

A:Cross-references: UNIPROT:Q04552; UNIPARC:UPI0000016BFBB; EMBL:Z29624; NID:g520879; PI

R;Cohen, M.B.; Schuler, M.A.; Berenbaum, M.R.
Proc. Natl. Acad. Sci. U.S.A. 89, 10920-10924, 1992
A;Title: A host-inducible cytochrome P-450 from a host-specific caterpillar: molecular cloning and characterization
A;Reference number: A46367; MUID:93066355; PMID:1279697
A;Accession: A46367
A;Status: preliminary
A;Molecule type: mRNA; protein
A;Residues: 1-23,'N',25-154,'NS',157-498 <COH>
A;Cross-references: UNIPARC:UPI0000126C97; GB:M80828; NID:g160763; PIDN:AAA29789.1; PID:
A;Note: sequence extracted from NCBI backbone (NCBIN:118719, NCBIP:118720)
C;Genetics:
A;introns: 445/1
C;Superfamily: human cytochrome P450 CYP3A5; cytochrome P450 homology
C;Keywords: chromoprotein; heme; iron; metalloprotein; oxidoreductase
F;300-465/Domain: cytochrome P450 homology <P45>
F;443/Binding site: heme iron (Cys) (axial ligand) #status predicted

Query Match 9.7%; Score 95; DB 1; Length 498;
Best Local Similarity 22.1%; Pred. No. 2;
Matches 43; Conservative 25; Mismatches 65; Indels 62; Gaps 6;
QY 36 DCEVCKEFLNRPYKSLIDRGVNSLDTIEKELISFCLD-----TKGENRLC 82
DB 84 DLDILHVLKIPKESPADRGVNSLDTIEKELISFCLD-----TKGENRLC 143
QY 83 YVLGATKDAATKILSEV--TRP--MSVHMPAMK-----ICEKLLKLD 121
DB 144 PLMSQVGDRIKIDVSTQTPQSIHNLVQKFTMTNIAACVFGNLDEGMLKTLEDL 203
QY 122 QICELKYKTKLDIASVDLRK-----MRVAELKQILHSMGEECRACA 162
DB 204 HIFTVNSAELDMYVPGILKLLNGSLFPKVVSKFFDNLTKNVLEMRKGTPSY----- 255
QY 163 EKTQYVNLQELAPK 177
DB 256 -QKQMDILQELREK 269

RESULT 3
BVRUC
excinuclease ABC, chain C - Escherichia coli (strain K-12)
N;Alternate names: excision nuclease ABC, chain C; uvrC protein
C;Species: Escherichia coli
C;Date: 31-Mar-1998 #sequence revision 23-Jan-1998 #text_change 01-Mar-2002
C;Accession: F64954; A22863; C24964; I78634
R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Cohen, M.;
A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A;Title: The complete genome sequence of Escherichia coli K-12.
A;Reference number: A64720; MUID:97426617; PMID:9278503
A;Accession: F64954
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-588 <BLAT>
A;Cross-references: UNIPARC:UPI00000D0580; GB:AE000284; GB:U00096; NID:g1788214; PIDN:AA
R;Gancar, G.B.; Sancar, A.; Rupp, W.D.
Nucleic Acids Res. 12, 4593-4608, 1984
A;Title: Sequences of the E. coli uvrC gene and protein.
A;Reference number: A22863; MUID:84247323; PMID:6330676
A;Accession: A22863
A;Molecule type: DNA
A;Residues: 1-268,'K',270-588 <SAN>
A;Cross-references: UNIPARC:UPI000016F603; GB:X03691; GB:X00189; GB:X00638; NID:g43287;
R;Sharma, S.; Stark, T.F.; Beattie, W.G.; Moses, R.E.
Nucleic Acids Res. 14, 2301-2318, 1986
A;Title: Multiple control elements for the uvrC gene unit of Escherichia coli.
A;Reference number: A93609; MUID:86176730; PMID:3515318
A;Accession: C24964
A;Molecule type: DNA
A;Residues: 1-96 <SHA>
A;Cross-references: UNIPARC:UPI00001748BD; GB:X03691; GB:X00189; GB:X00638; NID:g43287
R;Moolenaar, G.F.; van Sluis, C.A.; Backendorf, C.; van de Putte, P.

Nucleic Acids Res. 15, 4273-4289, 1987
A;Title: Regulation of the Escherichia coli excision repair gene uvrC. Overlap between
A;Reference number: A26750; MUID:87231005; PMID:3295776
A;Accession: I78634
A;Status: translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 'MSQFDKAFKTKVTSPQGVYR',1-6 <RES>
A;Cross-references: UNIPARC:UPI000016F604; EMBL:X05398; NID:g43291; PIDN:CAA28983.1; PI:
C;Genetics:
A;Gene: uvrC
A;Map position: 42 min
C;Function:
A;Description: one of the proteins involved in DNA excision repair, a process to remove
nuclease
C;Superfamily: excinuclease ABC chain C
C;Keywords: DNA repair

Query Match 9.1%; Score 89.5; DB 1; Length 588;
Best Local Similarity 25.1%; Pred. No. 7;
Matches 44; Conservative 29; Mismatches 51; Indels 51; Gaps 8;
QY 10 VAFCAGLLVSHPV-LTQGOEAGR-----PGAD--CEVCKEFLNRPY-----K 49
DB 235 VAFDAGMACVHVFIRQGVLSRSYFPKVPQGTSELSEVTFVGGFYLGSSQMRTPGE 294
QY 50 SLIDRGVNSLDTIEKELIS-----FCLDTGKENRLCVYLGATKDAATKILSEVTRPM 103
DB 295 ILDD--FNLSDKTLLADSLSELAGRKINVTQPRGDRARYLKLARTNAATLTSLKSQOS 352
QY 104 SVHMPAMKICEKLLKLDLSQICELKYKTKLDLASV-DLRKMRVAELKQILHSMGEE 157
DB 353 TVH-----QRLTALASVLKLPVKRMCECFDISHTMGEQ 385

RESULT 4
D85808
excinuclease ABC, chain C - Escherichia coli (strain O157:H7, substrain EDL933)
C;Species: Escherichia coli
C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 27-Nov-2001
C;Accession: D85808
R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhe
illier, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca
Nature 409, 529-533, 2001
A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A;Reference number: A85480; MUID:21074935; PMID:11206551
A;Accession: D85808
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-588 <STO>
A;Cross-references: UNIPARC:UPI00000D0580; GB:AE005174; NID:g12516011; PIDN:AAG56928.1;
A;Experimental source: strain O157:H7, substrain EDL933
C;Genetics:
A;Gene: uvrC
C;Superfamily: excinuclease ABC chain C

Query Match 9.1%; Score 89.5; DB 2; Length 588;
Best Local Similarity 25.1%; Pred. No. 7;
Matches 44; Conservative 29; Mismatches 51; Indels 51; Gaps 8;
QY 10 VAFCAGLLVSHPV-LTQGOEAGR-----PGAD--CEVCKEFLNRPY-----K 49
DB 235 VAFDAGMACVHVFIRQGVLSRSYFPKVPQGTSELSEVTFVGGFYLGSSQMRTPGE 294
QY 50 SLIDRGVNSLDTIEKELIS-----FCLDTGKENRLCVYLGATKDAATKILSEVTRPM 103
DB 295 ILDD--FNLSDKTLLADSLSELAGRKINVTQPRGDRARYLKLARTNAATLTSLKSQOS 352
QY 104 SVHMPAMKICEKLLKLDLSQICELKYKTKLDLASV-DLRKMRVAELKQILHSMGEE 157
DB 353 TVH-----QRLTALASVLKLPVKRMCECFDISHTMGEQ 385

RESULT 5

S13786
DNA-directed DNA polymerase (EC 2.7.7.7) III chain dnaX - Bacillus subtilis
N;Alternate names: DNA polymerase III (gamma and tau subunits) dnaX
C:Species: Bacillus subtilis
C:Date: 19-May-1994 #sequence revision 19-May-1994 #text_change 09-Jul-2004
C:Accession: S13786; S00745; S66049; B69618
R:Alonso, J.C.; Shiraheige, K.; Ogasawara, N.
Nucleic Acids Res. 18, 6771-6777, 1990
A:Title: Molecular cloning, genetic characterization and DNA sequence analysis of the
A:Reference number: S13786; MUID:91088245; PMID:2124672
A:Accession: S13786
A:Molecule type: DNA
A:Residues: 1-563 <ALO>
A:Cross-references: UNIPROT:P09122; UNIPARC:UPI0000005FPC3; EMBL:X17014; NID:G4532238; P:
R:Struck, J.C.R.; Vogel, D.W.; Ulbrich, N.; Erdmann, V.A.
Nucleic Acids Res. 16, 2720, 1988
A:Title: A dnaXZ-like open reading frame downstream from the Bacillus subtilis scRNA g
A:Reference number: S00745; MUID:88203213; PMID:2452406
A:Accession: S00745

A:Molecule type: DNA
A:Residues: 1-422 <STR>
A:Cross-references: UNIPARC:UPI000016E84A; EMBL:X06803; NID:G39891; PIDN:CAA29958.1; P: R:Ogasawara, N.; Nakai, S.; Yoshikawa, H.
DNA_Res.1, 1-14, 1994
A:Title: Systematic sequencing of the 180 kilobase region of the *Bacillus subtilis* chromosome
A:Reference number: S65967; MUID:96051385; PMID:7584024
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-563 <CGA>
A:Cross-references: UNIPARC:UPI000005FDC3; EMBL:D26185; NID:q467326; PIDN:BAAO5255.1; P: R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berte
C.; Bron, S.; Brouillet, S.; Bruchi, C.; Caldwell, B.; Capuano, V.; Carter, N.M.; Ch
A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E
Nature 390, 249-256, 1997
A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galle
techt, J.; Harwood, C.R.; Henauf, A.; Hilbert, H.; Holtsappel, S.; Hosono, S.; Hulio, M.P.
Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinio
A:Authors: Lauber, J.; Lazarevic, V.; Les, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maue
Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetel
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon
A:Authors: Schleicht, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Sero
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpestra, P.; Tognoni, A.; Tosato, V.; Uchiyama
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida,
A:Authors: Yoshikawa, H.F.; Zumstede, E.; Yoshikawa, H.; Danchin, A.
A:Title: The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*
A:Reference number: A69580; MUID:980404033; PMID:9184177

A;Accession: B69618
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-563 <KUN>
A;Cross-references: UNIPARC:UPI000005FDC3; GB:Z99104; GB:AL009136; NID:g2632267; PIDN:C
A;Experimental source: strain 168
C;Commentary

A:Gene: dnaX; dnaZX
A:Start codon: GTG
C:Superfamily: DNA-directed DNA polymerase III gamma chain
C:Keywords: nucleotidyltransferase

Query Match 9.1%; Score 89; DB 2; Length 563;
Best Local Similarity 20.8%; Pred. No. 7.4;
Matches 38; Conservative 32; Mismatches 55; Indels 58; Gaps 6;

Qy 39 VCKEFLNRPFKSLIDRGVNFSLDTIKELISFCLDYTKG-----ENRLCYV-----LGAT 88
Db 244 VSLQYIGKLAKSLDKKNVSDALETLNELL-----QQGKDPAPKLIEDMIFYFRDMLLYKT 297

Qy 89 KQAAATKILSEVT-----RPMISVHMPAMKICEKLKLLDS-----QI 123
Db 298 APGLGEGVKKYKVDYEFREI SEOTPRQALVEMTDILNKSHQPMKWTNEDDIPSEVAIVYKT 357

QY 124 CELKYEKTLDLASVDLRKMRVAELKQILHSWGEECRACAECTDYVNLIQELAPKYAATHP 183

Db 358 CQTSQSAADLPEVDMKMKIQLEQ-----EVERLKTGIIKAAAESP 400

Qy 184 KTE 186

Db 401 KKE 403

RESULT 8

hypotheetical protein F3E12.200 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C>Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 09-Jul-2004

C:Accession: T06133

R:Bevan, M.; Hilbert, H.; Braun, M.; Holzer, E.; Brandt, A.; Duesterhoeft, A.; Hoheisel, submitted to the Protein Sequence Database, April 1999

A:Reference number: Z15485

A:Accession: T06133

A:Molecule type: DNA

A:Residues: 1-828 <BEV>

A:Cross-references: UNIPROT:O65503; UNIPARC:UPI00000AB4B3; EMBL:AL022604; GSPDB:GN00062;

C:Genetics:

A:Gene: ATSP:F23E12.200

A:Map position: 4

A:Introns: 458/1; 505/3; 580/2

Query Match 9.0%; Score 88.5; DB 2; Length 828;

Best Local Similarity 22.9%; Pred. No. 13;

Matches 47; Conservative 26; Mismatches 65; Indels 67; Gaps 6;

Qy 19 SHPVLTCGGAGGAGRGADCEVCKEFLNRFYKS-----LIDRG----- 55

Db 404 SNATATRGCGGGGPRAPVEVAKEIENQFVKAESGSEIAKLEVGKHPYGRKHGTSSSA 463

Qy 56 ----VNFSLDITIEKLISFCLDTKGKRNLCYYLGATKDAATKILSEVTRPMSPVHPAMK 111

Db 464 AAADVPTVADIEELASRSNLSSTLKLHLW-----EKKLYHEVKAEEKLRLAHEK 516

Qy 112 ICEKLLKLSQICE-LKYET-----LDLASVDLRKMRVAEL----- 147

Db 517 KURKLLDQRGAEAIKVDKTRKLVDRMSTKIRIAIQVVDKISVTINKIRDEDLWQPLNA 576

Qy 148 -----KQILHWSGECRACAE 163

Db 577 LIQGLTRMKMTMLECHSQSQQAIRE 601

RESULT 9

arginine-rich protein - bovine (fragment)

C:Species: Bos primigenius taurus (cattle)

C>Date: 12-Feb-1998 #sequence_revision 13-Mar-1998 #text_change 10-Jul-1998

C:Accession: S69268

R:Galat, A.; Gerbod, M.C.; Bouet, F.; Riviere, S.

Arch. Biochem. Biophys. 330, 229-237, 1996

A:Title: Proteins and their amino acid compositions: uniqueness, variability, and applic

A:Reference number: S69268; MUID:96239137; PMID:8660651

A:Accession: S69268

A>Status: preliminary

A:Molecule type: protein

A:Residues: 1-39 <GAL>

A:Cross-references: UNIPARC:UPI000017C4A8

Query Match 9.0%; Score 88; DB 2; Length 39;

Best Local Similarity 57.4%; Pred. No. 0.43;

Matches 20; Conservative 3; Mismatches 12; Indels 0; Gaps 0;

Qy 36 DCBEVFLNRFYKSLIDRGVNFSLDTIEKELISF 70

Db 5 DEXVLSYLGRRFFQDLKDRDVTSPASIEKELIKF 39

RESULT 10

H81707

inclusion membrane localised protein Inca TC0396 [imported] - Chlamydia muridarum (stra

C:Species: Chlamydia muridarum, Chlamydia trachomatis MoPn

C>Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 09-Jul-2004

C:Accession: H81707

R:Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heideberg, J.F.; White, O.; Hickey, C.; Dodson, R.; Winn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg

Nucleic Acids Res. 28, 1397-1406, 2000

A:Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39.

A:Reference number: AB1500; MUID:20150255; PMID:10684935

A:Accession: H81707

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-276 <TET>

A:Cross-references: UNIPROT:Q9PKB8; UNIPARC:UPI00000578B9; GB:AE002306; GB:AE002160; NTU

C:Experimental source: strain Nigg (MoPn)

C:Genetics:

A:Gene: TC0396

Query Match 8.9%; Score 87.5; DB 2; Length 276;

Best Local Similarity 23.8%; Pred. No. 4.4;

Matches 36; Conservative 26; Mismatches 56; Indels 33; Gaps 5;

Qy 39 VCKEFLNRFYKSLID-----RGVNFSLDTIEKELISFCLDTKGKRNLCYYLGATKDAAT 93

Db 92 LCKTAPRLRLYKELQGEVASLKEVNFLLKSVQKE-----FLGLSKDFAT 134

Qy 94 --KILSEVTRP-----MSVHPAMKICRKLKLDLSQICELKYBKTLDLASVDLRKMR 143

Db 135 TSKDLSVDLSDFHNLQDFQSSHQGFEDLLEDFKNSAEDLRQIFSQETVQSLKSTLSLK 194

Qy 144 VAEKQILHWSGECRACAEKTDYVNLIQEL 174

Db 195 -BEIKEIVLPTREVRLAENKEDLLKIVQDL 224

RESULT 11

I40486

surfactin synthetase component II - Bacillus subtilis

N:Alternate names: surfactin synthetase srfA2; surfactin synthetase/competence protein

N:Contains: acid-amino-acid ligase (EC 6.3.2.-)

C:Species: Bacillus subtilis

C>Date: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 09-Jul-2004

R:Cosmina, P.; Rodriguez, F.; de Ferra, F.; Grandi, G.; Perego, M.; Venema, G.; van Sin

Mol. Microbiol. 8, 821-831, 1993

A:Title: Sequence and analysis of the genetic locus responsible for surfactin synthesis

A:Reference number: I40485; MUID:93360813; PMID:8355609

A:Accession: I40486

A>Status: translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-3587 <RES>

A:Cross-references: UNIPROT:Q04747; UNIPARC:UPI000005FF03; EMBL:X70356; MUID:9396480; PI

R:Hamoen, L.W.; Eshuis, H.; Jongbloed, J.; Venema, G.; van Sinderen, D.

Mol. Microbiol. 15, 55-63, 1995

A:Title: A small gene, designated comS, located within the coding region of the fourth

A:Reference number: S60866; MUID:95272393; PMID:7752896

A:Accession: S60866

A:Molecule type: DNA

A:Residues: 977-1104 <HAM>

A:Cross-references: UNIPARC:UPI00001787F1

R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berte

C: Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Ch

A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.

Nature 390, 249-256, 1997

A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galle

iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F

Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois

A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Mauee

Y.M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetell

Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon

A:Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Sero

akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terptrtra, P.; Tognoni, A.; Tosato, V.; Uchiyama, T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yanane, K.; Yasumoto, K.; Yata, K.; Yoshida, K.
A:Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.
A:Title: The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*.
A:Reference number: A69580; MUID:98044033; PMID:9384377
A:Accession: C69718
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-3587 <KUN>
A:Cross-references: UNIPARC:UPI000005PF03; GB:Z99105; GB:AL009126; NID:g2632457; PIDN:CA
A:Experimental source: strain 168
R:Fabre, C.; Quentin, Y.; Guisepi, A.; Busuttill, J.; Haele, J.; Denizot, F.
submitted to the EMBL Data Library, March 1993
A:Reference number: S46967
A:Accession: S46968
A:Molecule type: DNA
A:Residues: 1-32, 'F', 34-41, 'G', 43-109, 'D', 111-114, 'G', 116-138, 'V', 140-258, 'W', 260-308, 'A',
1758-1914, 'PK', 1917-2138, 'SRL', 2142, 'DSL', 2146-2444, 'Q', 2446-2712, 'H', 2714-2722, 'H', 272
A:Cross-references: UNIPARC:UPI00000B6366; EMBL:X72672; NID:G516358; PIDN:CAAS1223.1; PI
R:Fuma, S.; Fujishima, Y.; Corbelli, N.; D'Souza, C.; Nakano, M.M.; Zuber, P.; Yamane, K.
Nucleic Acids Res. 21, 93-97, 1993
A:Title: Nucleotide sequence of 5' portion of *srfA* that contains the region required for
A:Reference number: S35517; MUID:93181186; PMID:8441623
A:Accession: S35518
A:Status: significant sequence differences
A:Molecule type: DNA
A:Cross-references: EMBL:D13262; NID:g216345; PID:g216347
A:Experimental source: strain 168 trpC2
R:Borchert, S.; Patil, S.S.; Marahiel, M.A.
FEMS Microbiol. Lett. 92, 175-180, 1992
A:Title: Identification of putative multifunctional peptide synthetase genes using high
A:Reference number: S25658
A:Accession: S25658
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 168, 'C', 170-171, 281-283, 514-595, 597-647, 'R', 649-679, 'ETL', 683-693, 'DKR', 697,
A:Cross-references: UNIPARC:UPI00001787F2; UNIPARC:UPI00001787F3; EMBL:X65835; NID:g4020
A:Experimental source: strain ATCC 21332
C:Comment: This protein contains several amino acid-activating domains for the synthesis
the amino-terminal region of this protein, appear to be required for the development of
C:Genetics:
A:Gene: *srfAB*; *srfA2*
C:Superfamily: surfactin synthetase; acetate-CoA ligase homology; acyl carrier protein h
C:Keywords: antibiotic biosynthesis; carrier protein; duplication; ligase; phosphopant
F:511-951/Domain: acetate-CoA ligase homology <ACL1>
F:968-1035/Domain: acyl carrier protein homology <ACP1>
F:1036-1481/Domain: repeat <RP1>
F:1542-1995/Domain: acetate-CoA ligase homology <ACL2>
F:2013-2081/Domain: acyl carrier protein homology <ACP2>
F:2082-2529/Domain: repeat <RP2>
F:2591-3024/Domain: acetate-CoA ligase homology <ACL3>
F:3041-3108/Domain: acyl carrier protein homology <ACP3>
F:999,2045,3073/Binding site: phosphopantetheine (Ser) (covalent) #status predicted

Query Match 8.9%; Score 87.5; DB 2; Length 3587;
Best Local Similarity 24.7%; Pred. No. 81;
Matches 53; Conservative 23; Mismatches 70; Indels 69; Gaps 11;

Qy 3 CASPV-AVAFACGLL-----VSHVLTQGG-----EAGRPQ-----ADCEVC 40
Db 793 CLOPIGAPGELCVGGIGVARGVYNLPETLTKQFLEDPFGRITRYTGDRLARWLPDGNI- 851
Qy 41 KEFLNRYKSLIDRGVNFSLDTIEKEL-----ISFCLDTGKKNRLCYLIGATKDA 91
Db 852 -EFLGRIDNQVKGFRFELGEIETKLNMAEHVTEAAVIRKKNADENEICAVPTADREV 910
Qy 92 ATKILSEVTRPMSVHMPAMKICEKLLKLDQICELKYEKTLDLASVLRKMRVAELKQIL 151
Db 911 A-----VSLRKLTSQSLPDYVWPAHLIQWDS-----LPLTNGKINKKELP--- 952
Qy 152 HSWGEECRACAEKTDYVNLQELAPKYAATPKTE 186
Db 953 -----APQSEAQQ-----PEYAA--PKTE 969

RESULT 12

T10470

transcription initiation factor sigma 1 precursor, chloroplast - white mustard
C:Species: *Sinapis alba* (white mustard)
C>Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 09-Jul-2004
C:Accession: T10470
R:Kestermann, M.; Neukirchen, S.; Kloppstech, K.; Link, G.
Nucleic Acids Res. 26, 2747-2753, 1998
A:Title: Sequence and expression characteristics of a nuclear-encoded chloroplast sigma
A:Reference number: Z17035; MUID:98256449; PMID:9592164
A:Accession: T10470
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-481 <KES>
A:Cross-references: UNIPROT:O49935; UNIPARC:UPI000000ACE02; EMBL:Y15899; NID:g2706543; F
A:Experimental source: tissue type cotyledons
C:Genetics:
A:Gene: *Sig1*
A:Genome: nuclear
C:Superfamily: transcription initiation factor sigma mybA; transcription initiation fac
C:Keywords: chloroplast; DNA binding; sigma factor; transcription initiation

Query Match 8.9%; Score 87; DB 2; Length 481;
Best Local Similarity 23.2%; Pred. No. 9.1;
Matches 41; Conservative 35; Mismatches 61; Indels 40; Gaps 9;

Qy 33 PGADCEV-----CKEFLNRYKSLIDRGVNFSLDTIEKELISFCCLDTGKKNRLCYL 85
Db 90 PGSDQEELEDDIDHSVEALLLQRMLEKQWNLSEFKTRKKVPVTCGISARQR---I 145
Qy 86 GATKDAATKILSEVTRPMSV-----HMPAMKICEKLLK-----KLD---SQIC 124
Db 146 GAKKKTNVKAVSEVNFQNLKGVKGVISDHLVLSHAENVRLSKKIKSLGLRDLSEKSLT 205
Qy 125 E-LKYKTLDLASVLRKMRVAELKQILHSWGEECRACAEK--TDYVNLQELAPKY 178
Db 206 DRLCGPEDEQLAWSLKISR-AELQ-----AWLMECHLAREKLAWSNRLVNSIAQRY 257

RESULT 13

T09157

phosphoglucomutase precursor, chloroplast - spinach
C:Species: *Spinacia oleracea* (spinach)
C>Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 09-Jul-2004
C:Accession: T09157
R:Penger, A.; Pelzer-Reith, B.; Schnarrenberger, C.
Plant Physiol. 105, 1439-1440, 1994
A:Title: CDNA sequence for the plastidic phosphoglucomutase form *Spinacia oleracea* (L.)
A:Reference number: Z16596; MUID:95062723; PMID:7972501
A:Accession: T09157
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-583 <PEN>
A:Cross-references: UNIPROT:Q43156; UNIPARC:UPI000000ACA9D; EMBL:X75898; NID:g534981; PI
C:Genetics:
A:Gene: *pgm*
A:Genome: nuclear
C:Keywords: chloroplast
F:1-56/Domain: transit peptide (chloroplast) #status predicted <TNP>
F:57-583/Product: phosphoglucomutase #status predicted <MAT>

Query Match 8.9%; Score 87; DB 2; Length 583;
Best Local Similarity 26.4%; Pred. No. 11;
Matches 47; Conservative 25; Mismatches 56; Indels 50; Gaps 11;

Qy 16 LLVSHP---VLTOGQAGG-----RPGADCEVCCKEFLNRF-YKSLIDRGVNFSLDTIE 64
Db 334 VLKEHPETTIVTDARTSIGLSRFTNRGGKHC-----LYRVGYRNVIDKGVQLNEDDIE 387
Qy 65 KELISFCCLDTGK---KENRLCYLIGATKDAATKILSEVTRPMSVHMPAMKICEKLLKLD 120

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 13, 2005, 03:02:42 ; Search time 46 Seconds
(without alignments)
336.095 Million cell updates/sec

Title: US-10-648-361-2
Perfect score: 983
Sequence: 1 MWCASPVAVAFACAGLLVSH.....VNLQELAPKYAATHPKTEL 187

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.*
1: /cgn2_6/ptodata/1/iaa/5-COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/6-COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/H-COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/PCTUS-COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/RE-COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	101	10.3	507	2	US-09-091-097-10
2	101	10.3	507	2	US-10-109-670-12
3	93	9.5	507	2	US-09-091-097-34
4	93	9.5	507	2	US-10-109-670-34
5	90.5	9.2	1027	2	US-09-762-724-8
6	89.5	9.1	1029	2	US-09-762-724-6
7	89	9.1	322	2	US-08-489-039A-8442
8	89	9.1	498	1	US-08-457-274A-24
9	89	9.1	498	4	PCT-US95-05758-24
10	84.5	8.6	523	1	US-08-232-513A-3
11	83.5	8.5	450	2	US-10-104-047-2139
12	83	8.4	455	2	US-09-038-217A-18
13	83	8.4	455	2	US-09-447-034-18
14	83	8.4	463	1	US-08-336-408B-6
15	83	8.4	463	1	US-08-216-592A-8
16	83	8.4	463	4	PCT-US91-00399-6
17	82.5	8.4	617	2	US-09-489-039A-14321
18	81.5	8.3	472	2	US-09-166-350-17
19	80.5	8.2	963	2	US-09-384-272-12
20	80.5	8.2	963	2	US-09-394-272-13
21	80	8.1	719	2	US-09-641-741-28
22	80	8.1	1128	1	US-08-111-939-2
23	80	8.1	1128	2	US-09-641-741-30
24	80	8.1	1128	2	US-09-060-482-8
25	79.5	8.1	368	2	US-09-949-016-6301
26	79.5	8.1	373	2	US-09-949-016-7535
27	79.5	8.1	523	1	US-08-100-247-2

79.5 8.1 523 1 US-08-483-146A-2 Sequence 2, Appli
79.5 8.1 523 1 US-08-484-594A-2 Sequence 2, Appli
79.5 8.1 523 2 US-09-076-258A-2 Sequence 2, Appli
79.5 8.1 523 2 US-08-756-031-2 Sequence 2, Appli
79.5 8.1 533 2 US-09-248-796A-19201 Sequence 16, Appl
79 8.0 463 2 US-10-329-668-16 Sequence 16, Appl
79 8.0 466 2 US-09-949-016-7492 Sequence 1, Appli
79 8.0 524 2 US-09-352-548-1 Sequence 7492, Ap
79 8.0 524 2 US-09-949-016-6272 Sequence 6272, Ap
79 8.0 524 2 US-08-928-074-23 Sequence 23, Appl
79 8.0 535 2 US-09-949-016-8603 Sequence 8603, Ap
78.5 8.0 368 2 US-08-630-915A-20 Sequence 20, Appl
78.5 8.0 368 2 US-09-879-957-20 Sequence 20, Appl
78.5 8.0 473 2 US-09-716-964B-120 Sequence 120, App
78 7.9 734 2 US-03-438-185A-125 Sequence 125, App
78 7.9 1412 2 US-09-949-002-486 Sequence 486, App
78 7.9 1423 2 US-08-810-712-10 Sequence 10, Appl
78 7.9 1431 2 US-09-538-092-1198 Sequence 1198, Ap

ALIGNMENTS

RESULT 1
US-09-091-097-10
; Sequence 10, Application US/09091097
; Patent No. 6432407
; GENERAL INFORMATION:
; APPLICANT: TAKESAKO, KAZUTOH
; APPLICANT: OKADO, TAKASHI
; APPLICANT: YAGIHARA, TOMOKO
; APPLICANT: KURODA, MASANOBU
; APPLICANT: ONISHI, YOSHIMI
; APPLICANT: KATO, IKUNOSHIN
; APPLICANT: AKIYAMA, KAZUO
; APPLICANT: YASUEDA, HIROSHI
; APPLICANT: YAMAGUCHI, HIDEYO
; TITLE OF INVENTION: ANTIGENIC PROTEIN ORIGINATING IN
; TITLE OF INVENTION: MALASSEZIA
; NUMBER OF SEQUENCES: 58
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH, LLP
; STREET: PO BOX 747
; CITY: FALLS CHURCH
; STATE: VA
; COUNTRY: USA
; ZIP: 22040-0747
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/091.097
; FILING DATE:
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: WEINER, MARC S.
; REGISTRATION NUMBER: 32,181
; REFERENCE/DOCKET NUMBER: 1422-0346P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-205-8000
; TELEFAX: 703-205-8050
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 507 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-091-097-10

Query Match 10.3%; Score 101; DB 2; Length 507;
Best Local Similarity 26.6%; Pred. No. 0.012;


```
RESULT 5
US-09-762-724-8
; Sequence 8, Application US/09762724
; Patent No. 6664053
; GENERAL INFORMATION:
; APPLICANT: Kovacs, et al.
; TITLE OF INVENTION: Identification of a region of the major surface
; FILE REFERENCE: 4239-58054
; CURRENT FILING DATE: 2001-02-09
; PRIOR FILING DATE: 1999-08-17
; PRIOR FILING DATE: 1999-08-17
; PRIOR FILING DATE: 1998-08-17
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 1027
; TYPE: PRT
; ORGANISM: Pneumocystis carinii sp. f. hominis
US-09-762-724-8
Query Match          9.2%; Score 90.5; DB 2; Length 1027;
Best Local Similarity 25.9%; Pred. No. 0.48;
Matches 43; Conservative 24; Mismatches 56; Indels 43; Gaps 8;

Qy 34 GADCEVCKEFLNRFYKSLIDRGVNFSLDTIEKELISF-----CLDTKGKRNLCY- 83
Db 30 GLEDTKCKTLEEYCKTLTNAGLN--PEKVHEKLKDFCDNGKRNEKQDGLKNVKNQKCIK 87

Qy 84 YLGATKDAATKILSEVTRPMSVHPAMKICEKLLKLDISOICELKYEKTLDLASVDLRMR 143
Db 88 FOGKLOTAARKKISLTD-----EDCKKNEQCCUFLGACPTLKD-DCNKLJR 134

Qy 144 -----VAE--LKQILHSWGECRACAEKTDYVNLIOELAPK 177
Db 135 NNCYQKERNVAEBVLLRALRGDLNETKTEKK-----LKEVCPK 174

RESULT 6
US-09-762-724-6
; Sequence 6, Application US/09762724
; Patent No. 6664053
; GENERAL INFORMATION:
; APPLICANT: Kovacs, et al.
; TITLE OF INVENTION: Identification of a region of the major surface
; FILE REFERENCE: 4239-58054
; CURRENT FILING DATE: 2001-02-09
; PRIOR FILING DATE: 1999-08-17
; PRIOR FILING DATE: 1999-08-17
; PRIOR FILING DATE: 1998-08-17
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 1029
; TYPE: PRT
; ORGANISM: Pneumocystis carinii sp. f. hominis
US-09-762-724-6
Query Match          9.1%; Score 89.5; DB 2; Length 1029;
Best Local Similarity 25.9%; Pred. No. 0.63;
Matches 43; Conservative 24; Mismatches 56; Indels 43; Gaps 8;

Qy 34 GADCEVCKEFLNRFYKSLIDRGVNFSLDTIEKELISF-----CLDTKGKRNLCY- 83
Db 30 GLEDTKCKTLEEYCKTLTNAGLN--PEKVHEKLKDFCDNGKRNEKQDGLKNVKNQKCIK 87
```

```
Qy 84 YLGATKDAATKILSEVTRPMSVHPAMKICEKLLKLDISOICELKYEKTLDLASVDLRMR 143
Db 88 FOGKLOTAARKKISLTD-----EDCKKNEQCCUFLGACPTLKD-DCNKLJR 134

Qy 144 -----VAE--LKQILHSWGECRACAEKTDYVNLIOELAPK 177
Db 135 NNCYQKERNVAEBVLLRALRGDLNETKTEKK-----LKEVCPK 174

RESULT 7
US-09-489-039A-8442
; Sequence 8442, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 8442
; LENGTH: 322
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-8442
Query Match          9.1%; Score 89; DB 2; Length 322;
Best Local Similarity 25.1%; Pred. No. 0.15;
Matches 43; Conservative 21; Mismatches 71; Indels 36; Gaps 8;

Qy 2 WCASPVAVAFACAGLLVSHPVLTQOQAGRGPGADCEVCKEF-LNRFYKSLIDRGVNFSL 60
Db 75 WQYSPLNSIA-----VHALLTSGLEP-----QDVEIVVTPLTEFV----DEDAQVRL 119

Qy 61 DTIEKELISFCLDTKGKRNLCYVYGATKDAATKILSEVTRPMSVHPA-MKICEKLLKL 119
Db 120 DNIERKKSLRDVK-----LNKGVVFNITKVTVPESII--PAGISLDELKPS 166

Qy 120 DSQICELKYEKTLDLASVDLRKMRVAELKQILHSWGECRACAEKTDYVNL 170
Db 167 HSVLIIDLGTTLDISNV-----AGQMTSVSRIYGDPLKGLVSLTDAVKL 211
```

```
RESULT 8
US-08-457-274A-24
; Sequence 24, Application US/08457274A
; Patent No. 5734086
; GENERAL INFORMATION:
; APPLICANT: Scott, Jeffrey G.
; APPLICANT: Tomita, Takashi
; TITLE OF INVENTION: Cytochrome P450lpr Gene and Its Uses
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nixon, Hargrave, Devans & Doyle
; STREET: P.O. Box 1051, Clinton Square
; CITY: Rochester
; STATE: New York
; COUNTRY: USA
; ZIP: 14603
; COMPUTER READABLE FORM: disk
; MEDIUM TYPE: Floppy
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA: US/08/457,274A
; FILING DATE:
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Goldman, Michael L.
```

[illegible]

Patent No. 5700909
GENERAL INFORMATION:
APPLICANT: O'Brien, John S.
TITLE OF INVENTION: Proapsosin and Cytokine-Derived Peptides
TITLE OF INVENTION: as Therapeutic Agents
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: Campbell & Flores LLP
STREET: 4370 La Jolla Village Drive, Suite 700
CITY: San Diego
STATE: California
COUNTRY: USA
ZIP: 92122
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232,513A
FILING DATE: 21-APR-1994
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/100,247
FILING DATE: 30-JUL-1993
ATTORNEY/AGENT INFORMATION:
NAME: Campbell, Cathryn A.
REGISTRATION NUMBER: 31,815
REFERENCE/DOCKET NUMBER: P-UD 1643
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 535-9001

```

: PATENT NO. 3700909
:
: GENERAL INFORMATION:
:
: APPLICANT: O'Brien, John S.
: TITLE OF INVENTION: Proasoposin and Cytokine-Derived Peptides
: TITLE OF INVENTION: as Therapeutic Agents
:
: NUMBER OF SEQUENCES: 20
:
: CORRESPONDENCE ADDRESS:
: ADDRESS: Campbell & Flores LLP
: STREET: 4370 La Jolla Village Drive, Suite 700
: CITY: San Diego
: STATE: California
: COUNTRY: USA
: ZIP: 92122
:
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent In Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/232,513A
: FILING DATE: 21-APR-1994
: CLASSIFICATION: 514
:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/100,247
: FILING DATE: 30-JUL-1993
: ATTORNEY/AGENT INFORMATION:
: NAME: Campbell, Cathryn A.
: REGISTRATION NUMBER: 31,815
: REFERENCE/DOCKET NUMBER: P-UD 1643
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (619) 535-9001

```



```
;
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/216,592A
; FILING DATE: 23-MAR-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/825,667
; FILING DATE: 24-JAN-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: GOLDSTEIN, JORGE A
; REGISTRATION NUMBER: 29,021
; REFERENCE/DOCKET NUMBER: 1383.0060002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 463 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
;
US-08-216-592A-8

Query Match      8.4%; Score 83; DB 2; Length 463;
Best Local Similarity 21.3%; Pred. No. 1.1;
Matches 37; Conservative 21; Mismatches 60; Indels 56; Gaps 6;

QY 15 GLLVSHPLVTQGEAGGPPGA--DCEVCKEFLNRFYKSLIDRGVNFSLDTIEKELISFCL 72
Db 132 GSLVKHICAIIGDRSSGKHGYVSCGCKGFKR-----TIRKDLIYTC- 175
QY 73 DTGKENRLCYLGGATKDAATKILSEVTRPMSVHPAMKICEKLLKLDLSQICEKLYEKT 132
Db 176 ----RDNKDC-----LIDKQRNRCQYC--RYOKCL 200
QY 133 DLASVDLRKMRVAELKQILHSWGECRACAEKTDYVNLIQELAPKYAATHPKTE 186
Db 201 VMG---MKREAVQERQRERRAESEAEACASSHEDMPVERILEAEALAVEPKTE 251
```

Search completed: December 13, 2005, 03:11:29
Job time : 47 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 13, 2005, 02:55:21 ; Search time 187 seconds
(without alignment)

439.379 Million cell updates/sec

Title: US-10-648-361-2

Perfect score: 983

Sequence: 1 MWCASPVAVVAFAGLLVSH.....VNLQELAPKYAATHPKTEL 187

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_21:*

1: Geneseq1980s:*

2: Geneseq1990s:*

3: Geneseq2000s:*

4: Geneseq2001s:*

5: Geneseq2002s:*

6: Geneseq2003as:*

7: Geneseq2003bs:*

8: Geneseq2004s:*

9: Geneseq2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	983	100.0	187	5	ADG79397 Human sec
2	983	100.0	187	5	ADG79480 Human sec
3	983	100.0	187	6	AAO26450 Human chl
4	968	98.5	187	5	ADP69697 Human pol
5	968	98.5	187	8	ADMA4585 Novel hum
6	968	98.5	187	8	ADG98795 Protein f
7	894	86.9	163	6	AAO26451 Human chl
8	817	83.1	156	5	ADG79580 Human sec
9	546	55.5	179	6	ADG79580 Bovine pr
10	538	54.7	179	6	AAO26455 Human chl
11	538	54.7	179	9	ADW28228 Amino aci
12	538	54.7	182	8	ADQ96544 T cell ac
13	538	54.7	234	6	AAE34881 Human ARP
14	538	54.7	234	8	ADQ96546 T cell ac
15	538	54.7	234	8	ADG98166 Protein f
16	537	54.6	179	4	AAO262160 Mouse arg
17	537	54.6	179	6	ADG982552 Mouse pro
18	535	54.4	179	6	ADG982549 Human pro
19	535	54.4	234	6	ADG982548 Human arg
20	532.5	54.2	158	6	ADG982553 Mouse MAN
21	532.5	54.2	158	6	AAO26456 Human chl
22	532.5	54.2	159	6	ADG982554 A synthet
23	532.5	54.2	198	4	ADG982552 Amino aci
24	532	54.1	179	4	ADG982559 Human arg

25	532	54.1	234	4	ADG90765 Human she
26	532	54.1	234	8	ADMA4586 Human arg
27	532	54.1	234	8	ADN05973 Antipsori
28	532	54.1	234	8	ADO19303 Human PRO
29	532	54.1	234	8	ADP54366 Human PRO
30	532	54.1	234	8	ADP23376 PRO polyp
31	532	54.1	234	8	ADU06709 Novel bro
32	532	54.1	234	9	ADW28227 Amino aci
33	532	54.1	234	9	ADX06723 Cyclin-de
34	532	54.1	234	9	ADY14860 PRO polyp
35	529.5	53.9	158	6	ADG982550 Human MAN
36	529.5	53.9	159	6	ADG982551 A synthet
37	529	53.8	187	4	ADG65919 Amino aci
38	522.5	53.2	158	6	ADP58391 Human pol
39	518	52.7	179	6	ADG982559 Pig pro-M
40	495	50.4	94	6	AAO26452 Human chl
41	439	44.7	85	8	ADP29272 Human sec
42	419	42.6	106	4	ADG982552 Human rep
43	419	42.6	106	4	ADG982552 Human tes
44	410	41.7	94	8	ADG98190 Protein f
45	393	40.0	173	4	ADG98190 Protein f

ALIGNMENTS

RESULT 1
ADG79397

ID ADG79397 standard; protein; 187 AA.

XX ADG79397;

DT 11-MAR-2004 (first entry)

XX Human secreted protein of the invention SEQ ID NO:203.

XX ss; cytostatic; vasotropic; haemostatic; cardiovascular;

XX gastrointestinal; immunomodulator; inotropic; cerebroprotective;

XX neuroprotective; nephrotropic; antiinflammatory; antibacterial; virucide;

XX synaproteolical; antidiabetic; gene therapy; vaccine; cancer;

XX blood disorder; immune disorder; infection; inflammatory disorder;

XX type II diabetes; gene; human; secreted protein.

OS Homo sapiens.

XX WO200268638-A1.

XX 06-SEP-2002.

XX 21-FEB-2002; 2002WO-US005064.

XX 23-FEB-2001; 2001US-0270658P.

XX 12-JUL-2001; 2001US-0304444P.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Rosen CA, Komatsoulis GA, Birse CE, Choi GH, Olsen HS, Ni J;

XX Bell A;

XX WPI; 2002-750418/81.

XX N-PSDB; ADG79210.

XX New isolated polypeptide and encoding polynucleotide useful for

XX diagnosing, preventing, treating and/or ameliorating diseases such as

XX cancer, blood disorders, infections, inflammatory and immune disorders

XX and type II diabetes

XX Disclosure; SEQ ID NO 203; 336pp; English.

XX The invention relates to a novel isolated polypeptide. A protein of the

XX invention has cytostatic, vasotropic, haemostatic, cardiovascular,

XX gastrointestinal, immunomodulator, inotropic, cerebroprotective,

XX neuroprotective, nephrotropic, antiinflammatory, antibacterial, virucide,

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

CC gynaecological, and antidiabetic activity. A polynucleotide of the
 CC invention may have a use in gene therapy, and as a vaccine. The methods
 CC and compositions of the invention are useful for diagnosing, preventing,
 CC treating and/or ameliorating diseases such as cancer (neural,
 CC reproductive, gastrointestinal, endocrine, renal, CNS and respiratory
 CC neoplasias), blood disorders, immune disorders, infections, inflammatory
 CC disorders and type II diabetes. They can also be used in chromosome
 CC identification, screening assays and molecular weight markers. The
 CC present sequence is used in the exemplification of the invention.

XX Sequence 187 AA;

Query Match 100.0%; Score 983; DB 5; Length 187;
 Best Local Similarity 100.0%; Pred. No. 6.4e-95;
 Matches 187; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNCASPVAVAFACAGLLVSHPVLTQGEAGRGADCEVCCKEFLNRFYKSLIDRGVNFSL 60
 DB 1 MNCASPVAVAFACAGLLVSHPVLTQGEAGRGADCEVCCKEFLNRFYKSLIDRGVNFSL 60
 QY 61 DTIEKELISFCLDTKGKRNLCYILGATKDAATKILSEVTRPMSVHMPAMKICEKLLKLD 120
 DB 61 DTIEKELISFCLDTKGKRNLCYILGATKDAATKILSEVTRPMSVHMPAMKICEKLLKLD 120
 QY 121 SQICELKYKTKLDLASVDLRKMRVAELKQILHSWGEECRACAEKTDYVNLIOELAPKYAA 180
 DB 121 SQICELKYKTKLDLASVDLRKMRVAELKQILHSWGEECRACAEKTDYVNLIOELAPKYAA 180
 QY 181 THPKTEL 187
 DB 181 THPKTEL 187

RESULT 2

ADG79480
 ID ADG79480 standard; protein; 187 AA.

AC ADG79480;

XX 11-MAR-2004 (first entry)

DT Human secreted protein of the invention SEQ ID NO:286.

DE es; cytostatic; vasotropic; haemostatic; cardiovascular;
 KW gastrointestinal; immunomodulator; inotropic; cerebroprotective;
 KW neuroprotective; nephrotropic; antiinflammatory; antibacterial; virucide;
 KW gynaecological; antidiabetic; gene therapy; vaccine; cancer;
 KW blood disorder; immune disorder; infection; inflammatory disorder;
 KW type II diabetes; gene; human; secreted protein.

XX Homo sapiens.

XX WO200268638-A1.

XX 06-SEP-2002.

XX 21-FEB-2002; 2002WO-US005064.

XX 23-FEB-2001; 2001US-0270658P.

XX 12-JUL-2001; 2001US-0304444P.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Rosen CA, Komatsoulis GA, Bires CE, Choi GH, Olsen HS, Ni J;
 PI Bell A;

XX WPI; 2002-750418/81.

XX N-PSDB; ADG79293.

XX New isolated polypeptide and encoding polynucleotide useful for
 PT diagnosing, preventing, treating and/or ameliorating diseases such as
 PT cancer, blood disorders, infections, inflammatory and immune disorders
 PT and type II diabetes.

XX

XX Disclosure; SEQ ID NO 286; 936pp; English.

XX The invention relates to a novel isolated polypeptide. A protein of the
 CC invention has cytostatic, vasotropic, haemostatic, cardiovascular,
 CC gastrointestinal, immunomodulator, inotropic, cerebroprotective,
 CC neuroprotective, nephrotropic, antiinflammatory, antibacterial, virucide,
 CC gynaecological, and antidiabetic activity. A polynucleotide of the
 CC invention may have a use in gene therapy, and as a vaccine. The methods
 CC and compositions of the invention are useful for diagnosing, preventing,
 CC treating and/or ameliorating diseases such as cancer (neural,
 CC reproductive, gastrointestinal, endocrine, renal, CNS and respiratory
 CC neoplasias), blood disorders, immune disorders, infections, inflammatory
 CC disorders and type II diabetes. They can also be used in chromosome
 CC identification, screening assays and molecular weight markers. The
 CC present sequence is used in the exemplification of the invention.

XX Sequence 187 AA;

Query Match 100.0%; Score 983; DB 5; Length 187;
 Best Local Similarity 100.0%; Pred. No. 6.4e-95;
 Matches 187; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNCASPVAVAFACAGLLVSHPVLTQGEAGRGADCEVCCKEFLNRFYKSLIDRGVNFSL 60
 DB 1 MNCASPVAVAFACAGLLVSHPVLTQGEAGRGADCEVCCKEFLNRFYKSLIDRGVNFSL 60
 QY 61 DTIEKELISFCLDTKGKRNLCYILGATKDAATKILSEVTRPMSVHMPAMKICEKLLKLD 120
 DB 61 DTIEKELISFCLDTKGKRNLCYILGATKDAATKILSEVTRPMSVHMPAMKICEKLLKLD 120
 QY 121 SQICELKYKTKLDLASVDLRKMRVAELKQILHSWGEECRACAEKTDYVNLIOELAPKYAA 180
 DB 121 SQICELKYKTKLDLASVDLRKMRVAELKQILHSWGEECRACAEKTDYVNLIOELAPKYAA 180
 QY 181 THPKTEL 187
 DB 181 THPKTEL 187

RESULT 3

AAO26450

ID AAO26450 standard; protein; 187 AA.

XX AAO26450;

XX 07-FEB-2003 (first entry)

DT Human ch10-ARPR protein.

DE Cytostatic; chromosome 10 arginine-rich protein; ch10-ARPR; ARPR; cancer;

XX gene therapy; human.

XX Homo sapiens.

XX Key Location/Qualifiers

FT Peptide 1..24 /note= "Signal peptide"

FT Protein 25..187 /note= "Mature protein"

XX WO200279246-A2.

XX 10-OCT-2002.

XX 26-MAR-2002; 2002WO-EP003395.

XX 30-MAR-2001; 2001US-0280673P.

XX 24-MAY-2001; 2001US-0293453P.

XX (GENE-) GENEPROT INC.

XX Bougueleret L, Niknejad A, Bairoch A;

PI

XX WPI; 2003-040654/03.
DR N-PSDB; AAL53682.
XX
PT New isolated chromosome 10 arginine-rich protein related polypeptides,
PT useful for detecting and/or monitoring and treating conditions involving
PT aberrant expression of ARPR or uncontrolled growth of tissues, such as
PT cancer.
XX
PS Claim 1; Fig 1; 82pp; English.
XX
CC The invention relates to isolated chromosome 10 arginine-rich protein
CC related polypeptides (ch10-ARPR), or their fragments. The polypeptides,
CC nucleic acids and antibodies are useful for detecting and/or monitoring
CC and treating conditions involving aberrant expression of ARPR or
CC uncontrolled growth of tissues, such as cancer. The polypeptides are
CC useful as hybridisation probes, in chromosome and gene mapping, for the
CC generation of antisense RNA or DNA and in tissue or cell typing. The
CC methods are useful for detecting and measuring quantities of ARPR in
CC tissues and biological fluids. The host cells are useful for replicating
CC ARPR transcripts or expressing the ARPR proteins or polypeptides. The
CC polynucleotides of the invention can be used to treat disorders by gene
CC therapy. This sequence represents the human ch10-ARPR protein of the
CC invention
XX
SQ Sequence 187 AA;

Query Match 100.0%; Score 983; DB 6; Length 187;
Best Local Similarity 100.0%; Pred. No. 6.4e-95;
Matches 187; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MWCASPVAVAFACAGLLVSHPVLTQGEAGRGPCADCEVCKEFLNRFYKSLIDRGVNFSL 60
Db 1 MWCASPVAVAFACAGLLVSHPVLTQGEAGRGPCADCEVCKEFLNRFYKSLIDRGVNFSL 60

Qy 61 DTIEKELISFCLDTKGKCNRLCYVYGATKDAATKILSEVTRPMSVHPAMKICEKLLKLD 120
Db 61 DTIEKELISFCLDTKGKCNRLCYVYGATKDAATKILSEVTRPMSVHPAMKICEKLLKLD 120

Qy 121 SQICELKYKTLDLASVDLRKMRVAELKQILHSGEGECRACAEKTDYVNLIQELAPKYAA 180
Db 121 SQICELKYKTLDLASVDLRKMRVAELKQILHSGEGECRACAEKTDYVNLIQELAPKYAA 180

Qy 181 THPKTEL 187
Db 181 THPKTEL 187

RESULT 4
ABP69697
ID ABP69697 standard; protein; 187 AA.
XX
AC ABP69697;
XX
XX 20-JAN-2003 (first entry)
XX
XX Human polypeptide SEQ ID NO 1744.
XX
XX Human; genome mapping; gene therapy; food supplement; virus; fungus;
XX cell-proliferative disorder; neurodegenerative disease; bacterial;
XX Parkinson's disease; Alzheimer's disease; autoimmune disease;
XX multiple sclerosis; diabetes; genetic disorder; wound; burn; infection;
XX arthritis; cytostatic; immunomodulator; nontropic; neuroprotective;
XX antiparkinsonian; antidiabetic; immunosuppressive; dermatological;
XX haemostatic; vulnery; fungicide; antibacterial; virucide; protozoacide;
XX antiarthritic.
XX
XX Homo sapiens.
OS
XX
XX WO200270539-A2.
PN
XX
XX 12-SEP-2002.
PD
XX

PF 05-MAR-2002; 2002WO-US005095.
XX
PR 05-MAR-2001; 2001US-00799451.
XX
XX (HYSE-) HYSEQ INC.
PA
PI Tang YT, Zhou P, Goodrich RW, Asundi V, Zhang J, Zhao QA, Ren F;
PI Xue AJ, Yang Y, Ma Y, Yamazaki V, Chen R, Wang Z, Ghosh M,
PI Wehrman T, Wang J, Wang D, Drmanac RT;
XX
DR WPI; 2002-759812/82.
DR N-PSDB; ABZ11914.
XX
XX New polynucleotides comprising sequences assembled from expressed
PT sequence tags (ESTs), useful for treating cell-proliferative,
PT neurodegenerative, autoimmune, genetic, myeloid or lymphoid, or platelet
PT or coagulation disorders.
XX
PS Claim 9; SEQ ID NO 1744; 1012pp + Sequence Listing; English.
XX
CC The invention relates to an isolated polynucleotide (I) comprising a
CC nucleotide sequence selected from any of 948 sequences (ABZ1119-
CC ABZ12066) or their mature protein coding portion, active domain coding
CC protein or complementary sequences. The polynucleotides are useful for
CC identifying expressed genes or for physical mapping of human genome. The
CC encoded polypeptides (ABP68902-ABP68949) are useful as molecular weight
CC markers, as a food supplement, for generating antibodies, in medical
CC imaging, screening and diagnostic assays and for treating cell-
CC proliferative disorders (cancer), neurodegenerative diseases (Parkinson's
CC or Alzheimer's disease), autoimmune diseases (multiple sclerosis,
CC diabetes, lupus) genetic disorders, myeloid or lymphoid disorders,
CC platelet or coagulation disorders, wound, burns, incision, ulcers, liver
CC or lung fibrosis, infections (bacterial, viral, fungal, parasitic),
CC arthritis, etc. Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 187 AA;

Query Match 99.5%; Score 968; DB 5; Length 187;
Best Local Similarity 99.9%; Pred. No. 2.4e-93;
Matches 185; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MWCASPVAVAFACAGLLVSHPVLTQGEAGRGPCADCEVCKEFLNRFYKSLIDRGVNFSL 60
Db 1 MWCASPVAVAFACAGLLVSHPVLTQGEAGRGPCADCEVCKEFLNRFYKSLIDRGVNFSL 60

Qy 61 DTIEKELISFCLDTKGKCNRLCYVYGATKDAATKILSEVTRPMSVHPAMKICEKLLKLD 120
Db 61 DTIEKELISFCLDTKGKCNRLCYVYGATKDAATKILSEVTRPMSVHPAMKICEKLLKLD 120

Qy 121 SQICELKYKTLDLASVDLRKMRVAELKQILHSGEGECRACAEKTDYVNLIQELAPKYAA 180
Db 121 SQICELKYKTLDLASVDLRKMRVAELKQILHSGEGECRACAEKTDYVNLIQELAPKYAA 180

Qy 181 THPKTEL 187
Db 181 THPKTEL 187

RESULT 5
ADM44585
ID ADM44585 standard; protein; 187 AA.
XX
XX ADM44585;
AC
XX
XX 03-JUN-2004 (first entry)
DT
XX
XX Novel human arginine-rich protein-like polypeptide.
DE
XX
XX human; arginine-rich protein; cancer; inflammation; genetic disorder.
KW
XX
XX Homo sapiens.
OS

XX US2004053250-A1.
 XX 18-MAR-2004.
 XX 21-NOV-2002; 2002US-00302172.
 XX 05-MAR-2001; 2001US-00799451.
 XX 05-MAR-2002; 2002WO-US005095.
 XX 20-AUG-2002; 2002US-00225251.
 XX (TANG/) TANG Y T.
 XX (XUEA/) XUE A.
 XX (DRMA/) DRMANAC R T.
 XX Tang YT, Xue A, Drmanac RT;
 XX WPI; 2004-238579/22.
 XX N-PSDB; ADM44432.
 XX New isolated arginine-rich protein-like polynucleotides and polypeptides,
 XX useful for diagnosing and/or treating conditions associated with aberrant
 XX activity of the arginine-rich polypeptides, such as cancer and
 XX inflammation.
 XX Example 2; SEQ ID NO 949; 51pp; English.
 XX The invention relates to an isolated polynucleotide. The methods and
 XX compositions of the present invention are useful for the diagnosis and/or
 XX treatment of diseases or conditions associated with aberrant expression
 XX or activity of the arginine-rich protein-like polypeptides, such as
 XX cancer and inflammation. They can also be used in forensics, gene
 XX mapping, identification of mutations responsible for genetic disorders,
 XX and in assessing biodiversity. The present sequence represents a novel
 XX human arginine-rich protein-like polypeptide.
 XX Sequence 187 AA;

Query Match 98.5%; Score 968; DB 8; Length 187;
 Best Local Similarity 98.9%; Pred. No. 2.4e-93;
 Matches 185; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 MWCASPVAVVAFGAGLLVSHPVLTQGEAGRGADCEVCCKEFLNRFYKSLIDRGVNFSL 60
 Db 1 MWCASPVAVVAFGAGLLVSHPVLTQGEAGRGADCEVCCKEFLNRFYKSLIDRGVNFSL 60
 QY 61 DTIEKELISFCLDTKGENRLCYLGGATKDAATKILSEVTRPMSVHMPAMKICEKLKLD 120
 Db 61 DTIEKELISFCLDTKGENRLCYLGGATKDAATKILSEVTRPMSVHMPAMKICEKLKLD 120
 QY 121 SQICELKYEXTLDLASVDLRKMRVAELKQILHSWGECRACAEKTDYVNLIOELAPKYAA 180
 Db 121 SQICELKYEXTLDLASVDLRKMRVAELKQILHSWGECRACAEKTDYVNLIOELAPKYAA 180
 QY 181 THPKTEL 187
 Db 181 THPKTEL 187

RESULT 6
 ADS98795
 ID ADS98795 standard; protein; 187 AA.
 XX
 AC ADS98795;
 XX
 XX 30-DEC-2004 (first entry)
 XX Protein factor discovery related human contig polypeptide, SEQ ID 1059.
 XX antiinflammatory; cytostatic; antimicrobial; gene therapy; inflammation;
 XX leukaemia; nervous system disorder; infection.
 XX Homo sapiens.

XX WO2004087874-A2.
 XX 14-OCT-2004.
 XX 24-MAR-2004; 2004WO-US009202.
 XX 28-MAR-2003; 2003US-0458824P.
 XX (NUVE-) NUVELO INC.
 XX (DRMA/) DRMANAC R T.
 XX Tang YT, Zhou P, Wang J, Wang ZW, Hu T;
 XX WPI; 2004-737686/72.
 XX N-PSDB; ADS98455.
 XX New polynucleotides encoding a polypeptide with biological activity,
 XX useful for treating inflammation, leukemias, nervous system disorders, or
 XX infections.
 XX Example 3; SEQ ID NO 1059; 253pp; English.
 XX The invention relates to a novel isolated polynucleotide comprising any
 XX of the 235 nucleotide sequences described in the specification. The
 XX invention further comprises: an isolated polynucleotide encoding a
 XX polypeptide with biological activity, where the polynucleotide hybridizes
 XX to one of the 235 novel polynucleotides under stringent hybridization
 XX conditions, or having greater than about 99% sequence identity with the
 XX novel polynucleotide; a vector comprising a novel polynucleotide; an
 XX expression vector comprising the novel polynucleotide; a host cell
 XX genetically engineered to comprise the novel polynucleotide, which can be
 XX operatively associated with a regulatory sequence that modulates
 XX expression of the polynucleotide in the host cell; an isolated
 XX polypeptide encoded by the novel polynucleotide, or a polynucleotide
 XX hybridizing under stringent conditions to the novel polynucleotide; a
 XX composition comprising the polypeptide and a carrier; an antibody
 XX directed against the polypeptide; a method for detecting the novel
 XX polynucleotide in a sample; a method for detecting the polypeptide in a
 XX sample; a method for identifying a compound that binds to the polypeptide
 XX ; a method for producing the polypeptide; an isolated polypeptide
 XX comprising any of the 235 amino acid sequences described in the
 XX specification; and a collection of polynucleotides comprising of at least
 XX one of the polynucleotides cited above. The polypeptides and
 XX polynucleotides of the invention have antiinflammatory, cytostatic, and
 XX antimicrobial activities. The novel polynucleotide may be used to treat
 XX disorders by gene therapy. The polypeptides and polynucleotides are
 XX useful for treating inflammation, leukemias, nervous system disorders,
 XX or infections. This sequence represents the polypeptide encoded by
 XX contiguous DNA derived from one of the 235 novel isolated polynucleotides
 XX of the invention.
 XX Sequence 187 AA;

Query Match 98.5%; Score 968; DB 8; Length 187;
 Best Local Similarity 98.9%; Pred. No. 2.4e-93;
 Matches 185; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 MWCASPVAVVAFGAGLLVSHPVLTQGEAGRGADCEVCCKEFLNRFYKSLIDRGVNFSL 60
 Db 1 MWCASPVAVVAFGAGLLVSHPVLTQGEAGRGADCEVCCKEFLNRFYKSLIDRGVNFSL 60
 QY 61 DTIEKELISFCLDTKGENRLCYLGGATKDAATKILSEVTRPMSVHMPAMKICEKLKLD 120
 Db 61 DTIEKELISFCLDTKGENRLCYLGGATKDAATKILSEVTRPMSVHMPAMKICEKLKLD 120
 QY 121 SQICELKYEXTLDLASVDLRKMRVAELKQILHSWGECRACAEKTDYVNLIOELAPKYAA 180
 Db 121 SQICELKYEXTLDLASVDLRKMRVAELKQILHSWGECRACAEKTDYVNLIOELAPKYAA 180
 QY 181 THPKTEL 187
 Db 181 THPKTEL 187

RESULT 7
AAO26451
ID AAO26451 standard; protein; 163 AA.
XX AC AAO26451;
XX DT 07-FEB-2003 (first entry)
XX DE Human ch10-ARPR mature protein.
XX KW Cystostatic; chromosome 10 arginine-rich protein; ch10-ARPR; ARPR; cancer;
XX KW gene therapy; human.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
FT Modified-site 1
FT /note= "Residue is modified by pyrrolidone carboxylic
FT acid"
XX WO200279246-A2.
XX PN 10-OCT-2002.
XX PD 26-MAR-2002; 2002WO-EP003395.
XX PF 30-MAR-2001; 2001US-0280673P.
XX PR 24-MAY-2001; 2001US-0293453P.
XX PA (GENE-) GENEPROT INC.
XX PI Bougueleret L, Niknejad A, Bairoch A;
XX WPI; 2003-040654/03.
XX
XX New isolated chromosome 10 arginine-rich protein related polypeptides,
XX useful for detecting and/or monitoring and treating conditions involving
XX aberrant expression of ARPR or uncontrolled growth of tissues, such as
XX cancer.
XX
XX Claim 2; Fig 2; 82pp; English.
XX
XX The invention relates to isolated chromosome 10 arginine-rich protein
XX related polypeptides (ch10-ARPR), or their fragments. The polypeptides,
XX nucleic acids and antibodies are useful for detecting and/or monitoring
XX and treating conditions involving aberrant expression of ARPR or
XX uncontrolled growth of tissues, such as cancer. The polypeptides are
XX useful as hybridisation probes, in chromosome and gene mapping, for the
XX generation of antisense RNA or DNA and in tissue or cell typing. The
XX methods are useful for detecting and measuring quantities of ARPR in
XX tissues and biological fluids. The host cells are useful for replicating
XX ARPR transcripts or expressing the ARPR proteins or polypeptides. The
XX polynucleotides of the invention can be used to treat disorders by gene
XX therapy. This sequence represents the human ch10-ARPR mature protein of
XX the invention
XX
XX Sequence 163 AA;
XX
Query Match 86.9%; Score 854; DB 6; Length 163;
Best Local Similarity 100.0%; Pred. No. 2.1e-81;
Matches 163; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 25 QGQAGRGPGADCEVCCKEFLNRFYKSLIDRGVNFSLDTIEKELISFCLDTKGKRNLCYY 84
DB 1 QGQAGRGPGADCEVCCKEFLNRFYKSLIDRGVNFSLDTIEKELISFCLDTKGKRNLCYY 60
QY 85 LGATKQATKILSEVTRPMSVHPAMKICEKLLKLDLSQICELKYKTKLDLASVDLRQVR 144
DB 61 LGATKQATKILSEVTRPMSVHPAMKICEKLLKLDLSQICELKYKTKLDLASVDLRQVR 120
QY 145 AELKQILHSWGEBCRAEKTDYVNLQELAPKYAATHPKTEL 187

DB 121 AELKQILHSWGEBCRAEKTDYVNLQELAPKYAATHPKTEL 163
RESULT 8
ADG79580
ID ADG79580 standard; protein; 156 AA.
XX AC ADG79580;
XX DT 11-MAR-2004 (first entry)
XX DE Human secreted protein of the invention SEQ ID NO:386.
XX KW ss; cytostatic; vasotropic; haemostatic; cardiovascular;
XX KW gastrointestinal; immunomodulator; inotropic; cerebroprotective;
XX KW neuroprotective; nephrotropic; antiinflammatory; antibacterial; virucide;
XX KW gynaecological; antidiabetic; gene therapy; vaccine; cancer;
XX KW blood disorder; immune disorder; infection; inflammatory disorder;
XX KW type II diabetes; gene; human; secreted protein.
XX OS Homo sapiens.
XX PN WO200268638-A1.
XX PD 06-SEP-2002.
XX PF 21-FEB-2002; 2002WO-US005064.
XX PR 23-FEB-2001; 2001US-0270658P.
XX PR 12-JUL-2001; 2001US-030444P.
XX PA (HUMA-) HUMAN GENOME SCI INC.
XX PI Rosen CA, Komatsoulis GA, Birse CE, Choi GH, Olsen HS, Ni J;
XX PI Bell A;
XX WPI; 2002-750418/81.
XX
XX New isolated polypeptide and encoding polynucleotide useful for
XX diagnosing, preventing, treating and/or ameliorating diseases such as
XX cancer, blood disorders, infections, inflammatory and immune disorders
XX and type II diabetes.
XX
XX Disclosure; SEQ ID NO 387; 936pp; English.
XX
XX The invention relates to a novel isolated polypeptide. A protein of the
XX invention has cytostatic, vasotropic, haemostatic, cardiovascular,
XX gastrointestinal, immunomodulator, inotropic, cerebroprotective,
XX neuroprotective, nephrotropic, antiinflammatory, antibacterial, virucide,
XX gynaecological, and antidiabetic activity. A polynucleotide of the
XX invention may have a use in gene therapy, and as a vaccine. The methods
XX and compositions of the invention are useful for diagnosing, preventing,
XX treating and/or ameliorating diseases such as cancer (neural,
XX reproductive, gastrointestinal, endocrine, renal, CNS and respiratory
XX neoplasias), blood disorders, immune disorders, infections, inflammatory
XX disorders and type II diabetes. They can also be used in chromosome
XX identification, screening assays and molecular weight markers. The
XX present sequence is used in the exemplification of the invention.
XX
XX Sequence 156 AA;
XX
Query Match 83.1%; Score 817; DB 5; Length 156;
Best Local Similarity 100.0%; Pred. No. 1.5e-77;
Matches 156; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 32 RRGADCEVCCKEFLNRFYKSLIDRGVNFSLDTIEKELISFCLDTKGKRNLCYYLGATKDA 91
DB 1 RRGADCEVCCKEFLNRFYKSLIDRGVNFSLDTIEKELISFCLDTKGKRNLCYYLGATKDA 60
QY 92 ATKILSEVTRPMSVHPAMKICEKLLKLDLSQICELKYKTKLDLASVDLRQVRVAELKQIL 151
DB 61 ATKILSEVTRPMSVHPAMKICEKLLKLDLSQICELKYKTKLDLASVDLRQVRVAELKQIL 120

```

QY 152 HSWGECRACAEKTDYVNIQLAPKYAATHPKTEL 187
DB 121 HSWGECRACAEKTDYVNIQLAPKYAATHPKTEL 156

RESULT 9
ABB82558
ID ABB82558 standard; protein; 179 AA.
XX
AC ABB82558;
DT 05-FEB-2003 (first entry)
XX
DE Bovine pro-MANF polypeptide.
XX
KW Mature astrocyte-derived neurotrophic factor; MANF; neuroprotective;
KW dopaminergic; arginine-rich protein; ARP; pro-MANF; bovine.
XX
OS Bos sp.
XX
FH Key Location/Qualifiers
FT Peptide 1..21
FT /label= Signal_peptide
FT Protein 22..179
FT /label= Mature_protein
XX
PN WO200274956-A2.
XX
PD 26-SEP-2002.
XX
PF 20-MAR-2002; 2002WO-CA000373.
XX
PR 20-MAR-2001; 2001US-0277516P.
XX
PA (PRES-) PRESCIENT NEUROPHARMA INC.
XX
PI Commissiong JW, Raibekas AA;
XX
WPI; 2003-040555/03.
XX
New mature astrocyte-derived neurotrophic factor (MANF) polypeptide,
PT useful as a dopaminergic neuronal survival-promoting factor for treating
PT neurodegenerative diseases, e.g. Parkinson's disease or Alzheimer's
PT diseases.
XX
PS Disclosure; Fig 11A; 53pp; English.
XX
The invention relates to a substantially purified mature astrocyte-
CC derived neurotrophic factor (MANF) polypeptide. The MANF polypeptide is
CC useful for increasing survival of dopaminergic neurons, growing
CC dopaminergic neurons for transplantation, preventing dopaminergic
CC neuronal cell death in a mammal, or especially useful for treating a
CC disease or disorder of the nervous system in a patient. In particular,
CC the polypeptide is useful for treating neurodegenerative diseases, e.g.
CC Parkinson's disease, Alzheimer's diseases or amyotrophic lateral
CC sclerosis. The present sequence represents a bovine pro-MANF polypeptide
XX
SQ Sequence 179 AA;
Query Match 55.5%; Score 546; DB 6; Length 179;
Best Local Similarity 56.6%; Pred. No. 6.3e-49;
Matches 107; Conservative 28; Mismatches 42; Indels 12; Gaps 3;

QY 1 MWCASPVAVVAFAGLLVSHVLTQGEAGRPGADCEVCCKEFLNRYKSLIDRGVNFSL 60
DB 1 MWATHGLA-VALALSVPASRALRQ-----DCEVCISYLGRRFYQDLKDRDVTFSF 50
QY 61 DTTEKELISCLDTKGENRCLVYLGATKDAATKILSEVTRPMSVMPAMKICEKLLKLD 120
DB 51 ASIEKELIKFCRARGENRCLYIGATEDAATKILINEVSKRPLSHHLPVEKICEKLLKLD 110
QY 121 SQICELKYKTXLDLASVDLRMRVAELKQILHSWGECRACAEKTDYVNIQLAPKYA- 179

```

Db 1 MWATQGLA-VALALSVLPGRAL-----RPG-DCEVCISYLGRFYQDLKDRDVTSP 50
Qy 61 DTIEKELISFCLDTKGKRNLCYVYGATKDAATKILSVTRPMSVHMPAMKICEKLKKLD 120
Db 51 ATINENELIKFCREARGKRNLCYVYGATDDAATKIINEVSKPLAHHIPVEKICEKLKKD 110
Qy 121 SQICELKYKTDLDLASVDLRKMRVAELKQIILHSGEGRCAEKTDDYVNLIOELAPKYA- 179
Db 111 SQICELKYDKQIDLSTVDLKLRLVKELKKILDDGETCGCAEKSDYIRKINELMPKYAP 170
Qy 180 -ATHPKTEL 187
Db 171 KAASARTDL 179

RESULT 11

ADW28228
ID ADW28228 standard; protein; 179 AA.
XX AC
XX ADW28228;
XX DT 07-APR-2005 (first entry)
XX DE Amino acid sequence of human DG153 protein, shorter variant.
XX KW antidiabetic; anorectic; endocrine-Gen.; gene therapy; DG153; DG177;
XX KW pancreas disease; diabetes; obesity; metabolic syndrome;
XX KW metabolic disease; pancreatic regeneration; nutritional disorder.
XX OS Homo sapiens.

Key Location/Qualifiers

FT Misc-difference 10 /note= "Ala encoded by CGC"
FT Misc-difference 11 /note= "Leu encoded by GTC"
FT Misc-difference 176 /note= "Arg encoded by CCG"
XX WO2005005471-A2.

XX 20-JAN-2005.

XX 08-JUL-2004; 2004WO-EP007531.

XX 11-JUL-2003; 2003EP-00015883.

XX 22-JUL-2003; 2003EP-00016710.

XX (DEVE-) DEVELOPEN ENTWICKLUNGSBIOLOGISCHE FORSCH.

XX Onichtchouk D;

XX WPI; 2005-092064/10.

XX N-PSDB; ADW28226.

XX A composition for diagnosing, preventing or treating diabetes, obesity or
PT metabolic syndrome comprises a DG153 or DG177 polypeptide, nucleic acid
PT molecule or an effector of the polypeptide or nucleic acid molecule.

XX Claim 4; SEQ ID NO 3; 64pp; English.

XX The specification describes a pharmaceutical composition which comprises
CC a DG153 or DG177 protein, a nucleic acid molecule encoding a DG153 or
CC DG177 protein, and/or an effector/modulator of the nucleic acid molecule
CC or protein. The composition of the invention, and DG153 and DG177 nucleic
CC acid molecules and polypeptides, are useful for manufacturing an agent
CC for detecting, verifying, treating, alleviating or preventing pancreatic
CC diseases (e.g. diabetes such as insulin dependent diabetes mellitus or
CC non-insulin dependent diabetes mellitus), obesity, metabolic syndrome and
CC metabolic diseases or dysfunctions. The composition is also used for
CC manufacturing an agent for the modulation of pancreatic development or
CC for the regeneration of pancreatic tissue or cells. The present sequence

CC represents a human DG153 protein.

XX Sequence 179 AA;
SQ

Query Match 54.7%; Score 538; DB 9; Length 179;
Best Local Similarity 56.6%; Pred. No. 4.4e-48;
Matches 107; Conservative 26; Mismatches 44; Indels 12; Gaps 4;

Qy 1 MWASPVAVAFACGLIVSHVPLTQGEAGORPGADCEVCKEFLNRFYKSLIDRGVNFSL 60
Db 1 MWATQGLA-VALALSVLPGRAL-----RPG-DCEVCISYLGRFYQDLKDRDVTSP 50
Qy 61 DTIEKELISFCLDTKGKRNLCYVYGATKDAATKILSVTRPMSVHMPAMKICEKLKKLD 120
Db 51 ATINENELIKFCREARGKRNLCYVYGATDDAATKIINEVSKPLAHHIPVEKICEKLKKD 110
Qy 121 SQICELKYKTDLDLASVDLRKMRVAELKQIILHSGEGRCAEKTDDYVNLIOELAPKYA- 179
Db 111 SQICELKYDKQIDLSTVDLKLRLVKELKKILDDGETCGCAEKSDYIRKINELMPKYAP 170
Qy 180 -ATHPKTEL 187
Db 171 KAASARTDL 179

RESULT 12

ADQ96544
ID ADQ96544 standard; protein; 182 AA.
XX AC

XX ADQ96544;
XX AC

XX 07-OCT-2004 (first entry)

XX T cell activation associated protein #361.

XX antiallergic; antiarthritic; antiasthmatic; antidiabetic; anti-HIV;
XX antimicrobial; antirheumatic; immunosuppressive; neuroprotective;
XX gene therapy; T cell activation; diagnosis; autoimmune disease;
XX rheumatoid arthritis; asthma; multiple sclerosis; diabetes;
XX allergic disease; infectious disease; AIDS; chronic rejection; organ;
XX bone-marrow transplant.

XX Homo sapiens.

XX WO2004058805-A2.

XX 15-JUL-2004.

XX 25-DEC-2003; 2003WO-JP016715.

XX 26-DEC-2002; 2002JP-00376365.

XX 27-DEC-2002; 2002US-0436473P.

XX 25-APR-2003; 2003JP-00122113.

XX 28-APR-2003; 2003US-0465792P.

XX 21-OCT-2003; 2003JP-00360559.

XX 22-OCT-2003; 2003US-0512846P.

XX (ASAH-) ASahi Kasei PHARMA CORP.

XX Matsuda A, Yoneta S;

XX WPI; 2004-593134/57.

XX N-PSDB; ADQ96543.

XX New purified protein involved in T cell activation, useful for
PT diagnosing, preventing and/or treating acquired immunodeficiency
PT syndrome, autoimmune (e.g. rheumatoid arthritis, and diabetes), allergic
PT and infectious diseases.

XX Claim 1; SEQ ID NO 722; 2828pp; English.

XX The invention relates to purified proteins and genes encoding them, that
CC are involved in T cell activation (I) and has an amino acid deletion,

CC substitution or addition in the amino acid sequences. The methods and
 CC compositions of the present invention are useful for the diagnosis,
 CC prevention and/or treatment of autoimmune disease (rheumatoid arthritis,
 CC asthma, multiple sclerosis and diabetes), allergic disease, infectious
 CC disease, AIDS, and acute or chronic rejection at organ transplant or bone
 CC marrow transplant. This sequence corresponds to a protein involved in T
 CC cell activation.

XX
 SQ Sequence 182 AA;

Query Match 54.7%; Score 538; DB 8; Length 182;
 Best Local Similarity 56.6%; Pred. NO. 4.5e-48;
 Matches 107; Conservative 26; Mismatches 44; Indels 12; Gaps 4;

QY 1 MWCASPVAVVAFAGLLVSHPVLTQGEAGGPGADCEVCKEFLNRFYKSLDRGVNFSL 60
 DB 4 MMATQGLA-VALALSVLPGSRAL-----RPG-DCEVCISYLGFRFYQDLKDRDVTFSF 53
 QY 61 DTIEKELISFCIDTKGKRNLCVYLGATKDAATKILSEVTRPMSVHMPAMKICEKLKLD 120
 DB 54 ATTENELIKFCREARGKRNLCYIGATDDATKIINEVSKPLAHHPVEKICEKLKLD 113
 QY 121 SQICELKYBKTLDLASVDLRKMRVAELKQILHSWGEECRACAEKTDYVNLIOELAPKYA- 179
 DB 114 SQICELKYDKQIDLTSTVDLKLRLVKELKKILDWGGETCKGCAEKSDYIRKINELMPKYAP 173
 QY 180 -ATHPKTEL 187
 DB 174 KAASARTDL 182

RESULT 13

AAE34881
 ID AAE34881 standard; protein; 234 AA.

AC AAE34881;

XX 28-MAY-2003 (first entry)

DE Human ARP protein.

XX Behavioural disorder; attention deficit hyperactivity disorder; ADHD;
 KW Molecular marker; intellectual disorder; gene therapy; ARP; human.

XX Homo sapiens.

XX WO200290541-A1.

XX 14-NOV-2002.

XX 03-MAY-2002; 2002WO-AU000556.

XX 03-MAY-2001; 2001AU-00004756.

XX 04-JUN-2001; 2001AU-00005426.

XX 04-JUN-2001; 2001US-0295811P.

XX (MURD-) MURDOCH CHILDRENS RES INST.

XX (DELA/) DELATYCKI M.

XX Williamson R, Dahl HM, Forrest SM, Wilcox SA, De Silva MG;

XX Elliott KS, Lynch M;

XX WPI; 2003-111974/10.

XX New molecular marker of a behavioral disorder, useful for diagnosing
 PT behavioral disorder, or assessing the likelihood of developing behavioral
 PT disorder, e.g. Attention Deficit Hyperactivity Disorder or intellectual
 PT disorders.

XX Claim 21, Fig 11; 390pp; English.

XX The invention relates to a molecular marker of a behavioural disorder,
 CC which is in a genetic form, and comprises a genetic location on

CC chromosome 3 or an equivalent location on another chromosome, where a
 CC mutation at the location alone or in combination with environmental or
 CC other genetics factors is associated with or otherwise facilitates the
 CC development or progression of the behavioural disorder. The molecular
 CC marker is useful for diagnosing behavioural disorder, or assessing the
 CC likelihood of developing behavioural disorder, e.g. attention deficit
 CC hyperactivity disorder (ADHD) or intellectual disorders. They are also
 CC useful for facilitating the development of therapeutic protocols for
 CC treatment of the behavioural disorders. Sequences of the invention are
 CC useful in manufacturing a genetic probe to determine the likelihood of a
 CC subject having a behavioural disorder, such as ADHD. They are also useful
 CC for diagnosing, preventing or treating a behavioural disorder. The
 CC invention is useful in gene therapy. The present sequence is human ARP
 CC protein. This sequence is used in the invention

XX Sequence 234 AA;

Query Match 54.7%; Score 538; DB 6; Length 234;

Best Local Similarity 56.6%; Pred. NO. 6.2e-48;

Matches 107; Conservative 26; Mismatches 44; Indels 12; Gaps 4;

QY 1 MWCASPVAVVAFAGLLVSHPVLTQGEAGGPGADCEVCKEFLNRFYKSLDRGVNFSL 60

DB 56 MMATQGLA-VALALSVLPGSRAL-----RPG-DCEVCISYLGFRFYQDLKDRDVTFSF 105

QY 61 DTIEKELISFCIDTKGKRNLCVYLGATKDAATKILSEVTRPMSVHMPAMKICEKLKLD 120

DB 106 ATTENELIKFCREARGKRNLCYIGATDDATKIINEVSKPLAHHPVEKICEKLKLD 165

QY 121 SQICELKYBKTLDLASVDLRKMRVAELKQILHSWGEECRACAEKTDYVNLIOELAPKYA- 179

DB 166 SQICELKYDKQIDLTSTVDLKLRLVKELKKILDWGGETCKGCAEKSDYIRKINELMPKYAP 225

QY 180 -ATHPKTEL 187

DB 226 KAASARTDL 234

RESULT 14

ADQ96546

ID ADQ96546 standard; protein; 234 AA.

XX ADQ96546;

XX 07-OCT-2004 (first entry)

XX T cell activation associated protein #362.

DE antiallergic; antiarthritic; antiasthmatic; antidiabetic; anti-HIV;
 KW antimicrobial; antirheumatic; immunosuppressive; neuroprotective;
 KW gene therapy; T cell activation; diagnosis; autoimmune disease;
 KW rheumatoid arthritis; asthma; multiple sclerosis; diabetes;
 KW allergic disease; infectious disease; AIDS; chronic rejection; organ;
 KW bone-marrow transplant.

OS Homo sapiens.

XX WO2004058805-A2.

XX 15-JUL-2004.

XX 25-DEC-2003; 2003WO-JP016715.

XX 26-DEC-2002; 2002JP-00376365.

XX 27-DEC-2002; 2002US-0436473P.

XX 28-APR-2003; 2003JP-00121113.

XX 21-OCT-2003; 2003US-0465792P.

XX 22-OCT-2003; 2003JP-00360559.

XX (ASAH-) ASahi KASEI PHARMA CORP.

XX Matsuda A, Yoneta S;

XX WPI; 2004-593134/57.
 DR N-PSDB; ADQ96545.
 XX
 PT New purified protein involved in T cell activation, useful for
 PT diagnosing, preventing and/or treating acquired immunodeficiency
 PT syndrome, autoimmune (e.g. rheumatoid arthritis, and diabetes), allergic
 PT and infectious diseases.
 XX
 XX Claim 1; SEQ ID NO 724; 2828pp; English.
 PS
 CC The invention relates to purified proteins and genes encoding them, that
 CC are involved in T cell activation (i) and has an amino acid deletion,
 CC substitutions or addition in the amino acid sequences. The methods and
 CC compositions of the present invention are useful for the diagnosis,
 CC prevention and/or treatment of autoimmune disease (rheumatoid arthritis,
 CC asthma, multiple sclerosis and diabetes), allergic disease, infectious
 CC disease, AIDS, and acute or chronic rejection at organ transplant or bone
 CC -marrow transplant. This sequence corresponds to a protein involved in T
 CC cell activation.
 XX
 SQ Sequence 234 AA;
 Query Match 54.7%; Score 538; DB 8; Length 234;
 Best Local Similarity 56.6%; Pred. No. 6.2e-48;
 Matches 107; Conservative 26; Mismatches 44; Indels 12; Gaps 4;
 QY 1 MWCASPVAVVAFACGLLVSHPVLTQGEAGRGPCADCEVCKEFLNRRYKSLIDRGVNFSL 60
 DB 56 MWATQGLA-VALALSVLPGSRAL-----RPG-DCEVCISYLGFRYQDLKORDVTFSP 105
 QY 61 DTIEKELISFCLDTKGKRNLCYVILGATKDAATKILSEVTRPMSVHMPAMKICEKLKLD 120
 DB 106 ATIENELIKFCREARGKENRLCYIIGATDDAATKIINEVSKPLAHHPVEKICEKLKLD 165
 QY 121 SQICELKYKTLDLASVDLRMRVAELKQILHSGEECRACAEKTDYVNIQLAPKYA- 179
 DB 166 SQICELKYKQIDLSVDLKLKRVKELKILDDWGETCKGCAEKSDYIRKINELMPKYAP 225
 QY 180 -ATHPKTEL 187
 DB 226 KAASARTDL 234
 RESULT 15
 ADS98166
 ID ADS98166 standard; protein; 234 AA.
 XX
 AC ADS98166;
 XX
 XX 30-DEC-2004 (first entry)
 DT
 TX Protein factor discovery related isolated human polypeptide, SEQ ID 430.
 DE antiinflammatory; cytostatic; antimicrobial; gene therapy; inflammation;
 KW leukaemia; nervous system disorder; infection.
 XX
 XX Homo sapiens.
 OS
 XX
 PN WO2004087874-A2.
 XX
 PD 14-OCT-2004.
 XX
 PF 24-MAR-2004; 2004WO-US009202.
 XX
 PR 28-MAR-2003; 2003US-0458824P.
 XX
 PA (NUVE-) NUVELO INC.
 PA (DRMA/) DRMANAC R T.
 XX
 PI Tang YT, Zhou P, Wang J, Wang ZW, Hu T;
 XX WPI; 2004-737686/72.

DR N-PSDB; ADS97931.
 XX
 PT New polynucleotides encoding a polypeptide with biological activity, or
 PT useful for treating inflammation, leukemias, nervous system disorders, or
 PT infections.
 XX
 XX Claim 20; SEQ ID NO 430; 253pp; English.
 PS
 CC The invention relates to a novel isolated polynucleotide comprising any
 CC of the 235 nucleotide sequences described in the specification. The
 CC invention further comprises: an isolated polynucleotide encoding a
 CC polypeptide with biological activity, where the polynucleotide hybridizes
 CC to one of the 235 novel polynucleotides under stringent hybridization
 CC conditions, or having greater than about 99% sequence identity with the
 CC novel polynucleotide; a vector comprising a novel polynucleotide; an
 CC expression vector comprising the novel polynucleotide; a host cell
 CC genetically engineered to comprise the novel polynucleotide, which can be
 CC operatively associated with a regulatory sequence that modulates
 CC expression of the polynucleotide in the host cell; an isolated
 CC polypeptide encoded by the novel polynucleotide, or a polynucleotide
 CC hybridizing under stringent conditions to the novel polynucleotide; a
 CC composition comprising the polypeptide and a carrier; an antibody
 CC directed against the polypeptide; a method for detecting the novel
 CC polynucleotide in a sample; a method for detecting the polypeptide in a
 CC sample; a method for identifying a compound that binds to the polypeptide
 CC ; a method for producing the polypeptide; an isolated polypeptide
 CC comprising any of the 235 amino acid sequences described in the
 CC specification; and a collection of polynucleotides comprising of at least
 CC one of the polynucleotides cited above. The polypeptides and
 CC polynucleotides of the invention have antiinflammatory, cytostatic, and
 CC antimicrobial activities. The novel polynucleotide may be used to treat
 CC disorders by gene therapy. The polypeptides and polynucleotides are
 CC useful for treating inflammation, leukaemias, nervous system disorders,
 CC or infections. This sequence represents one of the 235 novel isolated
 CC polypeptides of the invention.
 XX
 SQ Sequence 234 AA;
 Query Match 54.7%; Score 538; DB 8; Length 234;
 Best Local Similarity 56.6%; Pred. No. 6.2e-48;
 Matches 107; Conservative 26; Mismatches 44; Indels 12; Gaps 4;
 QY 1 MWCASPVAVVAFACGLLVSHPVLTQGEAGRGPCADCEVCKEFLNRRYKSLIDRGVNFSL 60
 DB 56 MWATQGLA-VALALSVLPGSRAL-----RPG-DCEVCISYLGFRYQDLKORDVTFSP 105
 QY 61 DTIEKELISFCLDTKGKRNLCYVILGATKDAATKILSEVTRPMSVHMPAMKICEKLKLD 120
 DB 106 ATIENELIKFCREARGKENRLCYIIGATDDAATKIINEVSKPLAHHPVEKICEKLKLD 165
 QY 121 SQICELKYKTLDLASVDLRMRVAELKQILHSGEECRACAEKTDYVNIQLAPKYA- 179
 DB 166 SQICELKYKQIDLSVDLKLKRVKELKILDDWGETCKGCAEKSDYIRKINELMPKYAP 225
 QY 180 -ATHPKTEL 187
 DB 226 KAASARTDL 234

Search completed: December 13, 2005, 03:05:59
 Job time : 189 secs

This Page Blank (uspto)